

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Gatanaga, T.
Granger, G.A.
- (ii) TITLE OF INVENTION: Factors Altering Tumor Necrosis
Factor Receptor Releasing Enzyme Activity
- (iii) NUMBER OF SEQUENCES: 154
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: MORRISON & FOERSTER
 - (B) STREET: 755 PAGE MILL ROAD
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94304-1018
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: Windows
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0b
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: USSN 09/081,385
 - (B) FILING DATE: 014-NOV-1998
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME:
 - (B) REGISTRATION NUMBER:
 - (C) REFERENCE/DOCKET NUMBER: 22000-20577.21
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 650-813-5600
 - (B) TELEFAX: 650-494-0792
 - (C) TELEX: 706141

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4047 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAGCTTTTGG	CTTTCCTTCC	CCGGGAAAGG	CCGGGGCCAG	AGACCCGCAC	TCGGACCAGG	60
CGGGGGCTGC	GGGGCCAGAG	TGGGCTGGGG	AGGGCTGGGA	GGCGTCTGG	GGCCGGCTCC	120
TCCAGGCTGG	GGGCCCCAG	CTCCGGGAAG	GCAGTCTGG	CCTCGGATG	GGGCCGCGCG	180
TGGGGCCCGG	CGGGGCGGCC	TCGGGAGGCG	TCCAGGCTGC	GGGAGCGGGA	GGAGCGGCCG	240
TGCGGGCGCC	AGCGCCGTGG	GTGGAGGTCG	CCGTCCCTCC	TGAGGGGCAG	CCAGTGCGTT	300
TGGGACCCGG	GAGCAGAGCC	CGCGCTTCCC	CAGCGGCCTC	CCCGGGGGTC	TCACCGGGTC	360
ACCCGAGAGC	GGAGGCCCCG	GCTCCGCAGA	AACCCGGGGC	GGCCGCGGGG	AAGCAGCGCC	420
CTCAGGCGTC	GGAGGAGCCC	CCAGAAGGAC	CTCGCGCCTT	CCCGCCGGGC	TCCGACCGCC	480
TGGGTTCGGT	CGGGGACGGC	CCAGGCCGCG	AGGACCCCA	AGCGCAGCTC	AGTCTGCGGG	540
GCACGACCCA	GAGGCCAGCA	GCAGGAGACG	GGGCCGGGGC	CGGGAGAGGG	CGGGGAGGGC	600
GCTCCTGGGA	GGTCAAGGCC	AGGGCTAGAC	TTTCAGGGTC	ATGGCCTGGC	CCCTCATCCC	660
CAGGGAGGTG	AGGGGGCTCT	GTGAGCAGAG	GGGGCCCCGG	TGGAGAAGGC	GCTGCTAGCC	720

```

AGGGGCGGGG CAGGAGCCCA GGTGGGGACT TAAGGGTGGC TGAAGGGACC CTCAGGCTGC 780
AGGGATAGGG AGGGAAGCTA GGGGTGTGGC TTGGGGAGGT GCTGGGGGAC CGCGGGCGCC 840
CTTTATTCTG AAGCCGAATG TGCTGCCGGA GTCCCCAGTG ACCTAGAAAT CCAATTTCAAG 900
ATTTTCAGGA GTTTCAGGTG GAGACAAAGG CCAGGCCCCAG GTGAAAATGT GGCAGTGACA 960
GAGTATGGG TGAGAACCAC GGAGAGAGGA AGTCCCGAG GCGGATGATG GGACAGAGAG 1020
CGGGGACCAG AATTTTTTAA AACGCATCTG AGATGCGTTT GGCAGACTCA TAGTTGTTTT 1080
CCTTTACCG AGAAAGTGTG GGCAGAAAGC AGCTCTAAAG CCCAGGCTGC CCAGCCTGCA 1140
CTGGCAGAGC TGACGGAAGG CCAGGGCAGA GCCTTCCCTC CCTGTACAG ACATGAGCCC 1200
TGGAGATCTG GAATGAGGCA GATGTGCCA GGGAAAGCTG ATCCGCCCGG ACCCAGGGCC 1260
CCCCGGGTGC CCCTTTGAGC GTGGAATCGT TGCCAGGTCA TGGCTCCCTG CTATCGAACA 1320
CGGGACACGG GTCGTGTGCT GCACCTGGCA GTTGACAGGAC CGACACCCAC AATGCCTTAA 1380
GAGGTGATGA CTGCCTTCCA GGGGCTTGGC TGGCTGACAC TTTGCATGGC TCCTGGAGAA 1440
GAGGGATTGA GTGGAGTCCA CGGGTCATGG CCACGTCCTG GGTGCTGCCT CTGAGGCAGG 1500
GCCCCGCTGG GGTGAGAAGG GGTGAGAGAC AGGTTCCTGC CAGTTCAGCC TCTAACGGT 1560
GGTCTTCATG CCTAGGAACC CACTGGGGGC TTATGAAACT GCAGGTGGCT GAGTCTTGC 1620
CATGGGGTCT CTCTTCAGG AGGTCTGGG GGGGCCGGAG ACTGTACCC ACAAAGGGTC 1680
CCAGGTGAGG CGGATGTGGC CTGGCGCTGT GTGGCTCTGG ACCTAGTCTT TGGGCTTGG 1740
CTGGGCCCCA GGGCCTGGG TTGAGACAGC TGTGACGAG GCAAGCCATT TACCCCGTT 1800
TGGGGACAT TACATCTTC TAGCTTGGAA CACACAGGCA GCCAGGTTG TTATCCACAT 1860
TCCTCTCCA TGTCTTCTC TTGAGAACTT TTACCAGGTA TGTACGAGC TGGGCTCCAC 1920
CAGGGAGACT CAAGTGGAAA GCCCTCATCC TTGCTCTCCA GGAGACAGGA AAACCTATGG 1980
TTACAATTCC AGGGACAAGA GCGATGCATG TGAGGTGTGG CAAATCTCAC TGTCAACTG 2040
GAGAAATCAG AGACAGCTTC CTGGAGGAGC TGACACCTGG ACAGGCTTCT CCACAGGAGG 2100
AAGCGAGTGA GAGAAGCCAA CTGGGATGGA CCCATCATGT AGGGGGAACA GTGCGCGCAG 2160
AACCACAAC CACCCCAACC CTAGGCCAG AGCTCACGGA GAGAGCTGGG CCTCTCGGG 2220
TGACTACATA GTTCCCTGCT GGATCTTAGG TCTGTCTCTT GGGCAGCTCT GCTGAGACCT 2280
CTATGCGCTG TCCAGGCTGC ACCAAGGTTT TGTGACTATT GGTCTGGGT TGTTTGCGAG 2340
CAACTGAAGT GTTCTGTGT AAAACAGGCA CTTGATTGCT TGAAGGAAT GCTGTTGTT 2400
CTTGCTGCGA CAAACATTGA GCAGCATTTA GTGGGCGGT TATATCTGT GGAGTAATGG 2460
GTGTTTTTGA AGTCTGTCTT GGTACTGCA CATTAAAAGG AATATCATTT TCTGAAACAT 2520
TGCTATTTTC CACACCAGAA ATCATATCCT CTTGCTGGTC CATGTCTGAA GACCTTACAC 2580
GAGAAAGTCT TAATGTAAGT TTAGTAGAGT CCTGGATGG AGAACTAATT ATATCATACA 2640
TTGCCGCTT CTACTCTGC TCTTTTTCAT CTTGCTTAA TTTTATTTT TCTGCTTCT 2700
TTGTTTTTCT TTCTGGAGAA TCTAGCAAGA TATCTGGTGG AACATCTCGA GGTGATGAAC 2760
AAGGTAGAGA CTGAGATTGT AGGATTAAAG GTGGTCTTGA GCCTTTAGGA GTTCTTTCAC 2820
TTCCAGCAGG GGAGCATACT GGCTGTGGAG ATCTCAAGGG AAAAGATGCA GCATTCCTCA 2880
TTGTTGAAGA ATCTCCATCG TCACTACTTA GCCTGTGCAC CATGTGTAGG TAGTCTCAC 2940
TTGAACCATG TCTAGGATTA TCAGCATGAT GATTAGCTGA ATTGCCAGAC AACGGACCAG 3000
AAACTTTATT ATCATGTATG TTTCTCAAAC CACCTGCAAC AATGGGACTT GATACCGATG 3060
CTTGTTGCAT CTGTGGATGT GTTGTGTAA TGAAGGATG GGAATATGGC ATGTATCCTG 3120
CAGGGCTTTG TGGGGCGTAT GGACTAGGCA CTGGGCTATT TTGCTGTGGC ATAAATCTGT 3180
TCCAGAGCT TGTCTGTGGT GGCACAAACC GGCTGGAGGG GCTATGTGAG ATAGTGGTT 3240
GTTGATAATT GGAAGATGCA GGACTACTGT GCATGGAATT CTGAGAAAGT TTATACTGAG 3300
ACATCATCAT TCCACTTTGT ACATATCTGT TCTGCATGCT TTTCTCCCTG AAAACATTAG 3360
GACTCCTTGC CAGGACGGCC TGCAACAAGA CTGGTATGTC ACCTTCTGGG TCATCACTGC 3420
CAAGGTTATC TTTCAACTCT ATGTGATCTG TTGATACCTG GTTGAGGCTA TGGACAAGCT 3480
GTGAAACCAA ATTGTCATCT CTACAAGCCA AAAGGCAGT CACCTCTTCT GCTATTCTGT 3540
CATTAAAGAG AAGGCTCTTT GTAGTTGTAG CAGGTAAAGG AGATGGAAGA GGCAGCTGGT 3600
TCAGGAGGTC TGTGAGACTA GCAATCCCCG CAAGAGTAGT AATGGGGACA TGGGGCATAT 3660
CCCCATTCAAT CCTGAATTTT TGGAAATGGT TTGCCTATAA AAGTACTTAG TTCAGGTGCC 3720
AGCTGTCAAT ACTTCCATT TCCCAAACAC TGGGCGAATC GGCGTCTGAA TCCAAGGGGA 3780
GGCCGAGGCC GCTGTGGCGA GAGACTATAA TCCGGGCCGG GAGGGGGGGC GGCTACGGCT 3840
CCTCTCCGT CTCCTCAGT CGGGGAACAT GTAGAGCCGG GGGGAGACCA GCCGAGAAGA 3900
CAAATCGTTG CTTCTTCTC CTCCTCTCC TCCTTCTCCC ACATAGAAAC ACTCACAAAC 3960
ACCCGACCAC GGGCCCGAGC TACCGGGGGG GCATCGCCCG GGGCCCGGGA ACCAATTCTC 4020
CTGTGCGGG GGGCGTCCTT TGGATCC 4047

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 739 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

GGATCCAAAG GTCAAATCC CCACCTGGCA CTGTCCCCGG AGCGGGTCCG GCCCGGCCGG 60
CGCGCGGCCG GCGGCTTGGC GCCAGAAGCG AGAGCCCTC GGGGCTCGCC CCCCAGCCTC 120
ACCGGGTCAG TGAAAAAACG ATCAGAGTAG TGGTATTTCA CCGGCGGCCG GCAGGGCCCG 180
CGGACCCCGC CCGGGCCCCC TCGCGGGGAC ACCGGGGGGG CGCCGGGGGC CTCCCACTTA 240

```

TTCTACACCT	CTCATGTCTC	TTCACCGTGC	CAGACTAGAG	TCAAGCTCAA	CAGGGTCTTC	300
TTTCCCGCT	GATTCCGCCA	AGCCCGTTCC	CTTGGCTGTG	GTTCGCTGG	ATAGTAGGTA	360
GGGACAGTG	GAATCTCGTT	CATCCATTCA	TGCGCGTCAC	TAATTAGATG	ACGAGGCATT	420
TGGTACCTT	AAGAGAGTCA	TAGTTACTCC	CGCCGTTTAC	CCGCGCTTCA	TTGAATTTCT	480
TCACTTTGAC	ATTCAGAGCA	CTGGGCAGAA	ATCACATCGC	GTCAACACCC	GCCGCGGGCC	540
TTCGCGATGC	TTTGTTTTAA	TTAAACAGTC	GGATTCCCCT	GGTCCGCACC	AGTTCTAAGT	600
CGGCTGCTAG	GCGCCGGCCG	AAGCGAGGCG	CCGCGCGGAA	CCGCGGCCCC	CGGGGCGGAC	660
CCGCGGGGGG	GACCGGGCCG	CGGCCCTCC	GCCGCTGCC	GCCGCGCGCG	CCGCCGCGCG	720
CCGAAGAAGA	AGGGGGAAA					739

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CAAGAGTGGC	GGCCGACGCA	GGCCCCCGG	GTGCCCGGGC	CCCCCTCGAG	GGGGACAGTG	60
CCCCCGCCGC	GGGGGCCCGG	CGGCGGGCCG	CCGCCGGCCC	CTGCCGCCCC	GACCCCTCTC	120
CCCCCGCCGC	CGCCCCACG	CGGCGCTCCC	CCGGGGAGGG	GGGAGGACGG	GGAGCGGGGG	180
AGAGAGAGAG	AGAGAGAGGG	CGCGGGGTGG	CTCGTGCCGA	ATTCAAAAAG	CTT	233

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2998 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGATCCAAAG	AATTCCGCAC	GAGGTAGTCA	CGGCTCTTGT	CATTGTTGTA	CTTGACGTTG	60
AGGCTGGTGA	GCTTGGAAAA	GTCGATGCGC	AGCGTGCAGC	AGGCGTTGTA	GATGTTCTGC	120
CCGTCCAGCG	ACAGCTTGGC	GTGCTGGGCG	CTCACGGGGT	CCGCATACTG	CAGCAGGGCC	180
TGGAACCTGG	TGTTCTTGGT	GAAGGTGATG	ATCTTCAACA	CTGTGCCGAA	CTTGGAGAAA	240
ATCTGGTGCA	GCACATCCAG	GGTCACAGGG	TAGAAGAGGT	TCTCCACGAT	GATCCTGAGC	300
ACGGGGCTCT	GCCCGGCCAT	CGCCATCCCT	GCATCCACGG	CCGCCGCCGA	GGCAGCCAAG	360
GCCAGGTTCC	CCGACTGGAC	CGAGTTCACC	GCCTGCAGGG	CCGCCTGGGC	CCGCGCCTGG	420
TTGGGAGAGC	TGTCGGTCTT	CAGCTCCTTG	TGGTTGGAGA	ACTGGATGTA	GATGGGCTGG	480
CCGCGCAGCA	CAGGGGTAC	CGAGGTGTAG	TAGTTCACCA	TGGTATTGGC	AGCCTCCTCC	540
GTGTTTCATCT	CGATGAAGGC	CTGGTTTTTC	CCCTTCAGCA	TCAGGAGGTT	GGTGACCTTC	600
CCAAAGGGCA	GCCCCAGGGA	GATGACTTCC	CCCTCCGTGA	CGTCGATGGG	GAGCTTCCGG	660
ATGTGGATCA	CTCTAGAGGG	GACGCCTGCA	CTTCGGCTGT	CACCTTTGAA	CTTCTTGCTG	720
TCATTTCCGT	TTGCTGCAGA	AGCCGAGTTG	CTGCTCATGA	TAAACGGTCC	GTTAGTGACA	780
CAAGTAGAGA	AAAGCTCGTC	AGATCCCCGC	TTTGTACCAA	CGGCTATATC	TGGGACAATG	840
CCGTCCATGG	CACACAGAGC	AGACCCGCGG	GGGACGGAGT	GGAGGCGCCG	GAATCCTGGA	900
GCTAGAGCTG	CAGATTGAGT	TGCTGCGTGA	GACGAAGCGC	AAGTATGAGA	GTGCTCTGCA	960
GCTGGGCGCG	GCACTGACAG	CCCACCTCTA	CAGCCTGCTG	CAGACCCAGC	ATGCACTGGG	1020
TGATGCCTTT	GCTGACCTCA	GCCAGAAGTC	CCCAGAGCTT	CAGGAGGAAT	TTGGCTACAA	1080
TGCAGAGACA	CAGAACTAC	TATGCAAGAA	TGGGAAACG	CTGCTAGGAG	CCGTGAACCT	1140
CTTTGTCTCT	AGCATCAACA	CATTGGTCAC	CAAGACCATG	GAAGACACGC	TCATGACTGT	1200
GAAACAGTAT	GAGGCTGCCA	GGCTGGAATA	TGATGCCTAC	CGAACAGACT	TAGAGGAGCT	1260
GAGTCTAGGC	CCCCGGGATG	CAGGGACACG	TGGTCGACTT	GAGAGTGCCC	AGGCCACTTT	1320
CCAGGCCCAT	CGGGACAAGT	ATGAGAAGCT	GCGGGGAGAT	GTGGCCATCA	AGCTCAAGTT	1380
CCTGGAAAGAA	AACAAGATCA	AGGTGATGCA	CAAGCAGCTG	CTGCTCTTCC	ACAATGCTGT	1440
GTCCGCTTAC	TTTGCTGGGA	ACCAGAAACA	GCTGGAGCAG	ACCCTGCAGC	AGTTCAACAT	1500
CAAGCTGCGG	CCTCCAGGAG	CTGAGAAACC	CTCCTGGCTA	GAGGAGCAGT	GAGCTGCTCC	1560
CAGCCCAACT	TGGCTATCAA	GAAAGACATT	GGGAAGGGCA	GCCCCAGGGT	GTGGGAGATT	1620
GGACATGGTA	CATCCTTTGT	CACCTGGCCT	CTGGCTTGGG	CTCCTTTTTT	TGGCTGGGGC	1680
CTGACACCAG	TTTTGCCAC	ATTGCTATGG	TGGGAAGAGG	GCCTGGAGGC	CCAGAAGTTG	1740
CTGCCCTGTC	TATCTTCTG	GCCACAGGGC	TTCATTCCCA	GATCTTTTCC	TTCACCTTCA	1800
CAGCCCAACG	CTATGACAAA	ACCACTCCCT	GGCCAATGGC	ATCACTCTTC	AGGCTGGGGT	1860
GTGCTCCCTG	ACCAATGACA	GAGCCTGAAA	ATGCCCTGTC	AGCCAATGGC	AGCTCTTCTC	1920
GGACTCCCTT	GGGCCAATGA	TGTTGCGTCT	AATACCCTTT	GTCTCTCCTC	TATGCGTGCC	1980
CATTGCAGAG	AAGGGGACTG	GGACCAAAAG	GGTGGGGATA	ATGGGGAGCC	CCATTGCTGG	2040

GCCCACTGGG	GCCACCCAGG	GGCACCTGCA	CAGCCTGGGT	GCCATTGAAC	CAGTAGATCA	2700
GGCTGCTGTG	CTGGCTGTAG	TGCACCCAGA	GTCTGCTGTG	CCAGTTGGCA	TTGGGGCCAG	2760
GCATGGGCAT	GATCACCCTA	TCCCACGTGG	CAGCACCCCA	GAGTTTCCGC	AGCGCCCGCT	2820
CTGAGTAGTT	CTACCGGTCA	CAGCCCTTGG	CCACATTGGCT	GGTCTGCAGC	TCATATGGTG	2880
CCTGAATGTT	CCAGAGTGGT	TCATCACAGG	TTCTCAGGCG	GATACCCAGG	AACAAAGCCA	2940
AGCTCCCGAC	ACCTGGTGCA	TATTCGATCC	TTTTGTTCGA	GCCTTGCCAC	CTGGGTTCAT	3000
AGGTGGCGCT	CACCTGAATC	TCCACCTCAG	CATCATCTCA	TGCCCGCTTC	TGGCCACAGT	3060
CATAAGCTGT	CACTGTAAC	TTATAGAGCC	TCTCACCCT	GTA CTGCAGC	TTCTGTGTGT	3120
TTCCAATGTT	CCCGTCATTA	TCAATGAGGA	AAGGGGGTGT	GGGTGTGAGA	ATCTCATAGT	3180
AGCAGATCTG	CGCTGTACTG	GGGGAGACGT	CACCCGTCAA	GGCTTCCACC	CGCAGATGCG	3240
GATCGTACAG	CTTCCCCTCT	GTCACAGCCG	CACGATACAG	CCGTTCCACA	AACACTGGGG	3300
CAAACTCGTT	CAGCATCGTT	ACCCGCATAT	GCACAGTGGC	CTTGTGGGAC	TTCTTGGTGT	3360
TGGCCCGCTG	GGGGCCCTCG	CCACAGTCAT	AGBGCCTGAA	GGTAGAAGTG	TGTCCTTCTT	3420
GGGCCTCGCA	GTTCCACAGG	TCTTTGGCCC	GGATCAGCCC	CTCTCCTGTC	GCCTTGTCAA	3480
GGATCAGCAG	CTCAAAGGGC	ACCCCAAGCC	CATGGAGCCG	GGAAGCCGAC	ATCTCCATGT	3540
CATAGCGCAG	CGGGGACATC	TTGTTCAAAG	CAAAGAGTAG	TGAGTTCAGT	AGGACCGTGT	3600
TGTCATTCTC	GATACAGATG	CCCTGGTACT	CTGCCTCAAT	CCATGGCTTG	TGCTTGTGTG	3660
CTTTGTGTCA	GGAGCAGGAC	GCGCAGAGCAG	AGGCCAGTAG	AAGGGCGCAG	AGCAGAGAGG	3720
TCATGGTGGG	CGCTGGGGCA	GGGCAGGGCC	AGGCCTTTCG	CTCCCCGGG	AGCCTCCAGT	3780
CTCGCGATTC	CACCTTGGCG	GAGGGATACA	GGGGGGGAAA	ACCAAAATAA	AACGTCAAAT	3840
AAATTGTGTA	GGAGGAGTCC	AGCTTAGGAC	CGGGCCAGAG	CCAGGCCACG	CTCGGGGAGG	3900
GGGCCCTGTG	AGGTTACAGG	GATCACTGCT	GCCACCACCG	CCACCTCGGG	AGGCCATTAT	3960
TTTGCCATGG	CTCTTAGTGC	AACAGCTGCC	TCCTCTGTCA	TGGCAGACAG	CACCGTGATC	4020
AGGATCTCTT	CTCCACAGTC	GTA CTCTGCG	TCAATCTCCT	TGCCAAGGTC	TCCTCAGGG	4080
AGACGAAGGT	CTCTCTGTAC	CTCCCCGCTG	TCCTGGAGCA	GTGATAGGTA	CCCATCCTGG	4140
ATCTTTGGAT	CC					4152

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- 7) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3117 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGATCCAAAG	ATTGCGCACG	AGTGGCCACA	TCATGAACCT	CCAGGCCAG	CCCAAGGCTC	60
AGAACAAGCG	GAAAGCGTTC	CTCTTTGGGG	GCCAGGAACC	AGGCTCCAAAG	GAGCAGGCCC	120
CTCCCCTCGA	GGCCCCCAGC	CAGTCCATCA	GAGTGAAGGA	GGGACGATAC	CTCGGGCAGC	180
ATGGTCCAGG	AGGGGCGTTC	AGGCTCTCTC	AGCCTGTGGA	ACTGGCCCCC	CTCAGCAGCC	240
TGGCCCTGCT	GAACTCTGTG	GTGTATGGGC	CTGAGCGGAC	CTCAGCAGCC	ATGCTGTCCC	300
AGCAGGTGGC	CTCAGTAAGG	TGGCCCAACT	CTGTGATGGC	TCCAGGCGGG	GCGCCGGAGC	360
GTGAGAGGAG	TGGGGGTGTC	AGTGACAAGC	GCTGGGACGA	GACGCGAGCG	CAGCCTCCAC	420
CCCATTCAAC	ATGGAAGTGC	CACAGTCTGT	CCCTCTACAG	TGCAACCAAG	GGGAGCCCCG	480
ATCCTGGAGT	GGGAGTCCCG	ACTTACTATA	ACCACCTGAT	GGCAGTGAAG	CGGGAGAAAG	540
CGGGGGGCCC	ACAGCTGGAC	CGCTATGTGC	GACCAATGAG	GCCACAGAAG	GTGCAGCTGG	600
AGGTAGGGCG	GGCCCCAGCA	CCGTGTAATT	CTTTCCAGCG	AGCCAAGAAA	CCCCCAAACC	660
AGTCACTGCC	CCTGCAACCC	TTCAGCTGG	CATTGCGCCA	CCAGGTGAAC	CGGCGAGTCT	720
TCCGGCAGGG	CCACC CGGCC	CCAAAGCCGG	TGGCTGCCTT	CCCTCCACAG	AAGCAGCAGC	780
AGCAGAGCA	ACCACACAGC	CAGCAGCAGC	AGCAGCAGCG	AGCCCTACCC	CAGATGCCGC	840
TCTTTGAGAA	CTTCTATTCC	ATGCCACAGC	AACCTCGCA	GCAACCCAG	GACTTTGGCC	900
TGCAGCCAGC	TAGGCCACTG	AGGACAGTCC	ACCTGGCTCA	CCACAGCATG	GCACCTTACC	960
CTTCTCCCCC	CAACCCAGAT	GTAAACCCAG	AACCTGCGAA	GGCCCTTCTG	CAGGACTCAG	1020
CCCCCGAGCC	AGCGCTACCT	CAGGTCCAGA	TCCCCTTCCC	CCGCGCTCC	CGCCGCTCT	1080
CTAAGGAGGG	TATCTCGCTC	CCCAGCGCCC	TGGATGGGCG	TGGCAGCCAG	CCTGGGCGAG	1140
AGGGCACTGG	CACCTGTGTC	CTACATCACT	GGCCCTTGCA	GCGAGCCGCA	CTTGCTCTCC	1200
TGGGGCAGCC	CCATCTGAA	GCTCTGGGAT	TCCCGCTGGA	GTGAGGGGAG	TCGCGACTAC	1260
TGCCGTGATG	GGAGAGACTA	GCACCCAACT	GCCGGGAGCG	AGAGGCTCCT	GCCATGGGCA	1320
CGGAGGAGGG	CATGAGGGCA	GTGAGCAGCA	GGGAGCTGGG	GCAGGCTGCTA	CGGGGCGGAG	1380
TGATCCAGAG	CACGCGACGG	AGGCGCCGGG	CATCCGACGA	GGCCAAATTG	CTGACCTGG	1440
CCCAGAAGGC	TGTGGAGCTG	GCCTCACTGC	AGAATGCAA	GGATGGCAGT	GGTCTTGAAG	1500
AGAAGCGGAA	CAAGGTATTG	CGCTCAACTA	CCAAAGTGGG	GGTGGAGTGT	TCTGAGCCTT	1560
CTTCTAGGCA	AAGTCGAGCA	CGAGAGAGCA	TGGGATGTGT	ACCCCTCATC	ATCCGACTGT	1620
CTGTGCCTGT	GCGAACTGTG	GACCCAAGT	AGGCAGCCCA	GGCTGGAGGT	CTTGATGAGG	1680
ACGGGAAGGG	TCTTGAACAG	AACCTCGCTG	AGCACAAGCG	ATCAGTCATC	GTACACCCGA	1740
GGCGGTCCAG	CGGAATCCCC	GGGACAGATG	CTCAAGCTCA	GGCGAGGAGC	ATGAATGTCA	1800
AGTTTGAGGC	GGAGCCTTCC	GTGCGGAAAC	CAAGAGCAGC	GGCCAGGCCC	GAGCCCTCA	1860
TCATCCCCAC	CAAGGCGGGC	ACTTTCATCG	CCCTCCCGCT	CTACTCCAAC	ATGACCCCAT	1920
ACCAGAGCCA	CTCGGCTCT	CCCGTCGGCC	TAGCTGACCA	CCCCCTGAG	CGGAGCCTTG	1980
AGCTACTCTC	CTACACGGCG	CCCCCATCTC	TCCAGCCTGT	GGGGGAAGGC	TCTGGCCTCT	2040

ACTTCAATGC	CATCATATCA	ACCAGCACCA	TCCCTGCCCC	TCCTCCCATC	ACGCCTAAGA	2100
GTGCCCACG	CACGCTGCTC	CGGACTAACA	GTGCTGAAGT	AACCCCGCCT	GTCTCTCTCG	2160
TGATGGGGGA	GGCCACCCCA	GTGAGCATCG	AGCCACGGAT	CAACGTGGGG	TCCCGTTTCC	2220
AGGCAGAAAT	CCCCTTGATG	AGGGACCCTG	CCCTGGCAGC	TGCAGATCCC	CACAAGGCTG	2280
ACTTGGTGTG	GCAGCCATGG	GAGGACCTAG	AGAGCAGCCG	GGAGAAGCAG	AGGCAAGTGG	2340
AAGACCTGCT	GACAGCCGCG	TGCTCCAGCA	TTTTCCCTGG	TGCTGGCACC	AACCAGGAGC	2400
TGGCCCTGCA	CTGTCTGCAC	GAATCCAGAG	GAGACATCCT	GGAACGCTG	AATAAGCTGC	2460
TGCTGAAGAA	GCCCCTGCGG	CCCCACAACC	ATCCGCTGGC	AACCTATCAC	TACACAGGCT	2520
CTGACCAGTG	GAAGATGGCC	GAGAGGAAGC	TGTTCAACAA	AGGCATTGCC	ATCTACAAGA	2580
AGGATTTCTT	CCTGGTGACG	AAGCTGATCC	AGACCAAGAC	CGTGGCCAGT	TGCGTGGAGT	2640
TCTACTACAC	CTACAAGAAG	CAGGTGAAAA	TCGGCCGCAA	TGGGACTCTA	ACCTTTGGGG	2700
ATGTGGATAC	GAGCGATGAG	AAGTCGGCCC	AGGAAGAGGT	TGAAGTGGAT	ATTAAGACTT	2760
CCCAAAAGTT	CCCAAGGGTG	CCTCTTCCCA	GAAGAGAGTC	CCCAAGTGAA	GAGAGGCTGG	2820
AGCCCAAGAG	GAGAGTGAA	GAGCCAGGA	AGGAGGGGGA	GGAGGAGGTG	CCAGAGATCC	2880
AAGAGAAAGG	GGAGCAGGAA	GAGGGGCGAG	AGCGCAGCAG	GCGGGCAGCG	GCAGTCAAAG	2940
CCACGCAGAC	ACTACAGGCC	AATGAGTCGG	CCAGTGACAT	CCTCATCCTC	CGGAGCCACG	3000
AGTCCAACGC	CCCTGGGTCT	GCCGGTGGCC	AGGCCTCGGA	GAAGCCAAGG	GAAGGGACAG	3060
GGAAGTCACG	AAGGGCACTA	CCTTTTTCAG	AAAAAAAAAA	AAAAAACAA	AAAGCTT	3117

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3306 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAATTCGGCA	CGAGGTGAGT	TYCTGTGGA	ACACAGAGGC	TGCCTGTCCC	ATTCAGACAA	60
CGACGGATAC	AGACCAAGCT	TGCTCTATAA	GGGATCCCAA	CAGTGGATTT	GTGTTTAATC	120
TTAATCCGCT	AAACAGTTCG	CAAGGATATA	ACGTCTCTGG	CATTGGGAAG	ATTTTATGT	180
TTAATGTCTG	CGGCACAATG	CCTGTCTGTG	GGACCATCCT	GGGAAAACCT	GCTTCTGGCT	240
GTGAGGCAGA	AACCCAAACT	GAAGAGCTCA	AGAATTGGAA	GCCAGCAAGG	CCAGTCGGAA	300
TTGAGAAAAG	CCTCCAGCTG	TCCACAGAGG	GCTTCATCAC	TCTGACCTAC	AAAGGGCCTC	360
TCTCTGCCAA	AGGTACCGCT	GATGCTTTTA	TCGTCCGCTT	TGTTTGCAAT	GATGATGTTT	420
ACTCAGGGCC	CCTCAAATTC	CTGCATCAAG	ATATCGACTC	TGGGCAAGGG	ATCCGAAACA	480
CTTACTTTGA	GTTTGAAACC	GCGTTGGCCT	GTGTTCTCTC	TCCAGTGGAG	TGCCAAGTCA	540
CCGACCTGGC	TGGAAATGAG	TACGACCTGA	CTGGCCTAAG	CACAGTCAGG	AAACCTTGGA	600
CGGCTGTTGA	CACCTCTGTC	GATGGGAGAA	AGAGGACTTT	CTATTGAGC	GTGTTGCAATC	660
TCTTCCCTTA	CATTCTCTGA	TGCCAGGGCA	GCGCAGTGGG	GTCTTGCTTA	GTGTCAGAA	720
GCAATAGCTG	GAATCTGGGT	GTGGTGACAG	TGAGTCCCCA	AGCCCGGCGG	AATGGATCTT	780
TGAGCATCAT	GTATGTCAAC	GGTGACAAGT	GTGGGAACCA	GCGCTTCTCC	ACCAGGATCA	840
CGTTTGAGTG	TGCTCAGATA	TCGGGCTCAC	CAGCATTTC	GCTTCAGGAT	GGTTGTGAGT	900
ACGTGTTTAT	CTGGAGAAGT	GTGGAAAGCT	GTCCCGTTGT	CAGAGTGGAA	GGGGACAAGT	960
GTGAGGTGAA	AGACCCAAGG	CATGGCAACT	TGTATGACCT	GAAGCCCTCG	GGCCTCAACG	1020
ACACCATCGT	GAGCGCTGGC	GAATACACTT	ATTACTTCGG	GGTCTGTGGG	AAGCTTTCCT	1080
CAGACGCTCG	CCCCACAAGT	GACAAGTCCA	AGGTGGTCTC	CTCATGTGAG	GAAGGCGGG	1140
AACCGCAGGG	ATTTCAACAA	GTGGCAGGTC	TCCTGACTCA	GAAGCTAAGT	TATGAAATG	1200
GCTTGTTAAA	AATGAACCTC	ACGGGGGGGG	ACACTTGCCA	TAAGGTTTAT	CAGCGCTCCA	1260
CAGCCATCTT	CTTCTACTGT	GACCGCGGCA	CCCAGCGGCC	AGTATTTCTA	AAGGAGACTT	1320
CAGATTGTTC	CTACTTGTGT	GAGTGGCGAA	CGCAGTATGC	CTGCCACCTT	TTCGATCTGA	1380
CTGAATGTTT	ATTCAAAGAT	GGGGCTGGCA	ACTCCTTCGA	CCTCTCGTCC	CTGTCAAGGT	1440
ACAGTGACAA	CTGGGAAGCC	ATCACTGGGA	CGGGGGACCC	GGAGCACTAC	CTCATCAATG	1500
TCTGCAAGTC	TCTGGCCCGG	CAGGCTGGCA	CTGAGCCGTG	CCCTCCAGAA	GCAGCCGCGT	1560
GTCTGCTGGG	TGGCTCCAAG	CCCGTGAACC	TCGGCAGGGT	AAGGGACGGA	CCTCAGTGGA	1620
GAGATGGCAT	AATTGTCCTG	AAATACGTTG	ATGGCGACTT	ATGTCCAGAT	GGGATTCGGA	1680
AAAAGTCAAC	CACCATCCGA	TTACCTGCA	GCGAGAGCCA	AGTGAAGTCC	AGGCCCATGT	1740
TCATCAGCGC	CGTGGAGGAC	TGTGAGTACA	CCTTTGCCTG	GCCCAACAGC	ACAGCCTGTC	1800
CCATGAAGAG	CAACGAGCAT	GATGACTGCC	AGGTACACAA	CCCAAGCACA	GGACACCTGT	1860
TTGATCTGAG	CTCCTTAAGT	GGCAGGGCGG	GATTACAGC	TGCTTACAGC	GAGAAGGGGT	1920
TGGTTTACAT	GAGCATCTGT	GGGGAGAATG	AAAAGTGGCC	TCCTGGCGTG	GGGGCCTGCT	1980
TTGGACAGAC	CAGGATTAGC	GTGGGCAAGG	CCAACAAGAG	GCTGAGATAC	GTGGACCAAG	2040
TCCTGCAGCT	GGTGTACAAG	GATGGGTCCC	CTGTCCCTC	CAATCCGGGC	CTGAGCTATA	2100
AGAGTGTGAT	CAGTTTCGTG	TGCAGGCTG	AGGCCGGGCC	AACCAATAGG	CCCATGCTCA	2160
TCTCCCTGGA	CAAGCAGACA	TGCACCTCTC	TCTTCTCCTG	GCACACGCGG	CTGGCCTGCG	2220
AGCAAGCGAC	CGAATGTTCC	GTGAGGAATG	GAAGCTCTAT	TGTTGACTTG	TCTCCCTTAA	2280
TTCATCGCAC	TGGTGGTTAT	GAGGCTTATG	ATGAGAGTGA	GGATGATGCC	TCCGATACCA	2340
ACCTGATTTT	CTACATCAAT	ATTTGTGAGC	CACTAAATCC	CATGCACGGA	GTGCCCTGTC	2400
CTGCCGGAGC	CGCTGTGTGC	AAAGTTCCCTA	TTGATGGTCC	CCCCATAGAT	ATCGGCCGGG	2460
TAGCAGGACC	ACCAATACTC	AATCCAATAG	CAAATGAGAT	TACTTGAAT	TTTGAAAGCA	2520

GTACTCCTTG	CCAGGAATTC	AGTTGTAAT	AAAATTGAAC	CTGCTCAACA	GCTGAGGGAG	2580
ACTAGAAATG	ATGGGTCCAT	ATCCTGGTGC	ATTGTCATAC	AATTCAAACA	ATGGTGCAGC	2640
TACCAGCTTG	TAATTTTTAG	GGACTGCAAA	CAAGGCTTTT	TCTTGAACT	GAACCAGAAA	2700
CAACTTCTTA	TGTTCCITAG	GCTTTGTAAT	ATGTGCAGGA	ATATATGGAT	ACTGAGGAGG	2760
TTCAAAATTT	GGTCTCCACC	AGTTACCAAT	GCAATCGTCA	ATGACCCAGT	CTTGCAAAAC	2820
TCCATCCTGA	CGACCCAGTA	TCTCTGTCT	TAAGCGTTTT	AGTCCTTCAA	CTTCATCTTC	2880
TCCTGGGTGA	AGTTCACCAC	CAGGTAGTTT	GAAGAAAGTT	GTTCCAGCT	GCAGCAGTAA	2940
CACATGGGGT	AGCCGGTGCT	CATGTACAAT	CAGAACCCTT	TCTACAGTCC	TCCTCATTCC	3000
AATTTTATCA	AATCTTCCCT	TCATGCGCTG	AAATCTGGCT	GCAACAGAGC	TGTCCTTCTC	3060
GTAGAGGGGC	TCITTTGTAC	CAAAAGTATA	ATTGGTAAGA	GGGTACAGGT	TGATGGTGCG	3120
CTCCAGGGTG	AGGGGCTTCG	TCTGCTGGAT	GTAATTGTTG	CCGAAGTGA	TGACCCCGCG	3180
GGGCCAGCCG	GTCTGCGAGC	GATTGGGCGG	TACCACAGAC	ATGCTGGCGA	GCTCCGGCGC	3240
TGACGGCGAG	CAGAAAGTGG	CAGGCAGGGT	AGACTTTCCC	CGTGGCGGAA	GCCTCGTGCC	3300
GAATTC						3306

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4218 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAATTCGGCA	CGAGAATGGA	TCAACCTCAA	CAACACGTTA	AAGCTAGACG	AAAGAAGTAA	60
TACACAGTGT	ATGAGTCTCA	CATGAAATAC	CCGGATGTAA	ATCCAAGGAA	ACAGGAAGCA	120
GATTGGTGGT	TGCCAGGGAC	AAGGGCGGTG	GGAGGAGAAA	ATGGAGAGTA	ACGGGACTTT	180
ACTTTTGGAG	TGATGAGAAT	GTTTTGGAGC	TAGATAGAAG	TGGTGGTTGT	ACACCATTTG	240
GGATGTACTA	CCACTTAATT	GTTCACTTAA	AAAGTTAATT	TATGTGAATT	GCATCTTAAT	300
TAAAAACAAG	GATAACATTC	CAACTCCTGG	ACATTATCCT	TCCTTTCCAT	TTGATGTCAG	360
GCCCGTGTTA	GAATTCTCAT	CCGGTTTGGT	CACTGCACCT	AAGATGTGGA	GAAATTAGGA	420
CGCACAGTTA	AGAGGAAGGA	TAACACTGAT	TAAGGTAGTG	CTTTTCTAGG	TTTCCCCTAA	480
ACAATTTAAC	AGATGGATAG	TGGCACCCT	TACGAGATGG	AAAAACCAGC	GGAAGGAAGA	540
TTTGGGGGAG	AAGTTAAGTT	TGTCTTGGGC	CTGTGTTTTG	CAACCTGAGT	GTAAGAGACA	600
TATGTAAAGT	CTTCAGTGGC	GAACACTTAA	AACTAGAAAT	GGATCAGAAT	TTTATCTTTG	660
GATGTGACTT	CTCAAGGATG	GTCTTGTGAC	TTCAGTGCCT	GGTCAAAATGA	CAAGATGGGC	720
AATCTTTTCC	TGAAGGTCCA	AGCACCTGAA	CGTGGCAGGG	TGACCCGATT	CCGATTTGCT	780
TAGAACAATC	CTAGTTTCATG	CCTATTGTCC	CTCATGTAAT	TAATATCACT	CTCAAAATGT	840
CTCATTTTGT	GCAATAAATT	CTGCAACGTG	ATGGCGCGAC	TCTCGCGGCC	CGAGCGGCCG	900
GACCTTGTCT	TCGAGGAAGA	GGACCTCCCC	TATGAGGAGG	AAATCATGCG	GAACCAATTC	960
TCTGTCAAAT	GCTGGCTTCA	CTACATCGAG	TTCAAACAGG	GCGCCCCGAA	GCCAGGCTC	1020
AATCAGCTAT	ACGAGCGGGC	ACTCAAGCTG	CTGCCCTGCA	GCTACAAACT	CTGGTACCGA	1080
TACCTGAAGG	CGCGTCCGGC	ACAGGTGAAG	CATCGCTGTG	TGACCGACCC	TGCCATGAA	1140
GATGTCAACA	ACTGTCATGA	GAGGGCCTTT	GTGTTTCATG	ACAAGATGCC	TCGCTCTGTG	1200
CTAGATTACT	GCCAGTTCTT	CATGGACCAG	GGGCGCGTCA	CACACACCCG	CCGCACCTTC	1260
GACCGTGCCC	TCCGGGCACT	GCCCATCACG	CAGCACTCTC	GAATTTGGCC	CCTGTATCTG	1320
CGCTTCTCTG	GCTCACACCC	ACTGCCTGAG	ACAGCTGTGC	GAGGCTATCG	GCGCTTCTCT	1380
AAGCTGAGTC	CTGAGAGTGC	AGAGGAGTAC	ATTGAGTACC	TCAAGTCAAG	TGACCGGCTG	1440
GATGAGGCGC	CCCAAGCGCT	GGCCACCGTG	GTGAACGACG	AGCGTTTCGT	GTCTAAGGCC	1500
GGCAAGTCCA	ACTACCAGCT	GTGGCAGGAG	CTGTGCGACC	TCATCTCCCA	GAATCCGGAC	1560
AAGGTACAGT	CCCTCAATGT	GGACGCCATC	ATCCGCGGGG	GCCTCACCCG	CTTCACCGAC	1620
CAGCTGGGCA	AGCTCTGGTG	TTCTCTCGCC	GACTACTACA	TCCGCAGCGG	CCATTTGAG	1680
AAGGCTCGGG	ACGTGTACGA	GGAGGCCATC	CGGACAGTGA	TGACCGTGCG	GGACTTCACA	1740
CAGGTGTTTG	ACAGCTACGC	CCAGTTTCAG	GAGAGCATGA	TCGCTGCAAA	GATGGAGACC	1800
GCCTCGGAGC	TGGGGCGCGA	GGAGGAGGAT	GATGTGGACC	TGGAGCTGCG	CCTGGGCCGC	1860
TTTCGAGCAGC	TCATCAGCCG	GCGGCCCTTG	CTCCTCAACA	GCGTCTTGCT	GCGCCAAAAC	1920
CCACACCACG	TGCACGAGTG	GCACAAGCGT	GTCGCCCTGC	ACCAGGGCCG	CCCCGGGAG	1980
ATCATCAACA	CCTACACAGA	GGCTGTGCAG	ACGGTGGACC	CCTTCAAGGC	CACAGGCAAG	2040
CCCCCACTC	TGTGGGTGGC	GTTTGCCAA	TTTTATGAGG	ACAACGGACA	GCTGGACGAT	2100
GCCCGTGTC	TCCTGGAGAA	GGCCACCAAG	GTGAACCTCA	AGCAGGTGGA	TGACCTGGCA	2160
AGCGTGTGGT	GTCAGTGGCG	AGAGCTGGAG	CTCCGACACG	AGAACTACGA	TGAGGCCTTG	2220
CGGCTGCTGC	GAAAGGCCAC	GGCGCTGCCT	GCCCGCCGGG	CCGAGTACTT	TGATGGTTCA	2280
GAGCCCGTGC	AGAACCGCGT	GTACAAGTCA	CTGAAGGTCT	GGTCCATGCT	CGCCGACCTG	2340
GAGGAGAGCC	TGGCACCTTT	CCAGTCCACC	AAGGCCGTGT	ACGACCGCAT	CCTGGACCTG	2400
CGTATCGCAA	CACCCAGAT	CGTCATCAAC	TATGCCATGT	TCCTGGAGGA	GCACAAGTAC	2460
TTTCGAGGAG	GCTTCAAGGC	GTACGAGCGC	GGCATCTCGC	TGTTCAAGTG	GCCCAACGTG	2520
TCCGACATCT	GGAGCACCTA	CCTGACCAAA	TTCATTGCCC	GCTATGGGGG	CCGCAAGCTG	2580
GAGCGGGCAC	GGGACCTGTT	TGAACAGGCT	CTGGACGGCT	GCCCCCAAAA	ATATGCCAAG	2640
ACCTTGATCC	TGCTGTACGC	ACAGCTGGAG	GAGGAGTGGG	GCCTGGCCCG	GCATGCCATG	2700
GCCGTGTACG	AGCATGCCAC	CAGGGCCGTG	GAGCCCGCCC	AGCAGTATGA	CATGTTCAAC	2760

ATCTACATCA	AGCGGGCGGC	CGAGATCTAT	GGGGTACACC	ACACCCGCGG	CATCTACCAG	2820
AAGGCCATTG	AGGTGCTGTC	GGACGAGCAC	GGCGTGAGA	TGTGCTGCG	GTTTGCAGAC	2880
ATGAGTGACA	ACCTCGGGGA	GATTGACCGC	GCCCGGGGCA	TCTACAGACT	GTGCTCCCAG	2940
ATTCTGTGACC	CGCGGACGAC	GGCGCGTGTC	TGGCAGAGCT	GGAAGAGCTT	TGAGGTCGCG	3000
CATGGCAATG	AGGACACCAT	CAAGGAAATG	CTGCGTATCC	GGCGCAGCGT	GCAGGCCACG	3060
TACAACACGC	AGGTCAACTT	CATGGCCTCG	CAGATGCTCA	AGGTTCTCGG	CAGTGGCCACG	3120
GGCACCGTGT	CTGACCTGGC	CCCTGGGCAG	AGTGGCATGG	ACGACATGAA	CGCTGTGGAA	3180
CAGCGGGCAG	AGCAGCTGGC	GCTGTGAGCG	GAGCGTGACC	AGCCCTTGCG	CGCCGAGAGC	3240
AAGATCCTGT	TCGTGAGGAG	TGACGCGCTCC	CGGGAGGAGC	TGGCAGAGCT	GGCACAGCAG	3300
GTCAACCCCG	AGGAGATCCA	GCTGGGCGAG	GACGAGGACG	AGGACGAGAT	GGACCTGGAG	3360
CCCCAACGAG	TTCGGCTGGA	GACGACAGCA	GTCGACGCGC	CAGTGTGTTG	GAGCGTGAAG	3420
GAAGACTGAC	CCGTCCCCTC	GTGCCGAATT	CGGCACGAGC	AAGACCAGCC	CCCAGATCAT	3480
TTGCCTCAAA	GGTTTTCCCT	CGAAGTCACA	AATGTTTCAA	GGAATCTCAA	ATTTTACAAA	3540
GTTTGAAGTG	TGGGACTTGG	TGGCCTGTGG	CTGTGTCTCT	TCTCTGTAGC	TGTTTTCTCC	3600
CTACATCCCT	GAAAGGAAGT	TGGGCTGTGC	CTCTCATCCG	CAGACCTCCC	TTTCCAGCGC	3660
CCAGGGCATG	GGGTGCTGTG	AGGGCAGCAG	GCTAGGTGTG	ACCGTGTCTC	TGGCCTCCAG	3720
GCCCGGTGTC	CTCTGTCTCT	TAGCCCACTA	AGGCCCTGGC	CCATTGTGTC	TAAACAGGCA	3780
TCGGGACCTA	GAAAGAGACA	ACAATCTCTC	TGGGTACGCA	CTCTGGCTAG	GACGTGGTCT	3840
CCTGACTGGG	ATCCAGGCCT	TCTCCCCTCG	CCATGTGAAT	TCCAGGGGCG	AGAGCCTGAA	3900
ATTGTGAACA	CAGCATGGC	CAAGAGATGA	TACCCGTGGG	AACCGAGGCT	CTCTTCTCTC	3960
CTTGCTGCTG	TTCTGGGTTG	CAGAGTAGTG	GAGGCTGTCT	TGAGAGGAGT	TGGAGTGCTG	4020
GTTTTCACCC	TGGTGGTGTT	GTTTGTCTTT	GAGGGACATT	AGAAAGCCCA	CGCCAGCCCT	4080
TGCTCCTGGC	CTGCACACAG	CGGAGCGACT	TTTCTAGGTA	TGCTCTTGAT	TTCTGCAGAA	4140
GAGCAGCGGT	GTGATGGAGCC	AAGAGGAAGT	GTGACTGAAA	CTGTCCACTC	ATAGCCCGGC	4200
TGCCGTATTG	AGAGGGCT					4218

GAGCTCGCGC	GCCTGCAGGT	CGACACTAGT	GGATCCAAG	AATTCGGCAC	GAGGAAACT	60
CAACCGGTGTA	CGAGTGGAGG	ACAGGCGACG	AGCCCTCTGT	GGTGGAAACG	CCCCACTAGG	120
AGGAGCTTCT	TGAGCAGGTG	GCAGAAAGAT	CGATTGACTG	GGGCGACTTT	GGGGTAGCTG	180
CAGTGTCTGA	GGGGACTGAC	TCTGGCATCT	TGTCGGAGGC	TGCTGGAATC	GACTGGGGCA	240
TCTTCCCGGA	ATCAGATTCA	AAGGATCCTG	GAGGTGATGG	GATAGACTGG	GGAGACGATG	300
CTGTTTGCTT	GCAGATCACA	GTCGTGGAAG	CAGGAACCCA	GGCTCCAGAA	GGTGTGTCCA	360
GGGGGCGTGA	TGCCCTCTGC	CTGCTTGAAT	ACACTGAGAC	CCGGAATCAG	TTCCTTGATG	420
AGCTCATGGA	GCTTGAGATC	TCTTtagccc	AGAGAGCAGT	GGAGTTGAGT	GAGGAGGCAG	480
ATGTCCTGTG	TGTGAGCCAG	TTCAGCTGTG	CTCGGCAACT	TGCGCAGGGC	ACAGCCAAAG	540
AAGAAGTGGT	TACCATTTGT	GAGTGTCTGG	AGGATCTGAT	TGGCAAGCTT	CACAGCTCTT	600
AGCTGCAACA	CCTGTTTATG	ATCTTGCCCT	CACCAAGGTA	TGTGGACCGA	GTGACTGAAT	660
TCCTTCAAGCA	AAAGCTGAAG	CAGTCCCAAG	TGCTGCTTT	GAAGAAAGAG	CTGATGGTGC	720
AAGAAGCAGCA	GGAGGCACCT	GAGGAGCAGG	CGGCTCTGGA	GCCTAAGCTG	GACCTGCTAC	780
TGGAGAAAGAC	CAGGAGACTT	CAGAAGCTGA	TTGAAGCTGA	CATCTCCAAG	AGGTACAGCT	840
GGCGCCCTGT	GAACTTGATG	GGAACCTTCT	TGTGACACCT	TCCGTGTTCT	TGCTTGCCCA	900
TCTTTCTCCG	TTTTGGGATG	AAGATGATCA	CCAGGGCTGT	TGTTTtGGGG	CCCTTCAAGG	960
CAAAAGACCA	GGCTGACTGG	AAGATGGAAG	GGCAAGAGAA	GGAAAGCGCA	CTGTATGGTG	1020
ATCTTGGCAC	TCTCATGTT	CTCTACAAGA	AGCTGTGGTG	ATTGGCCCTG	TGCTTATCA	1080
GCGGAAAACCT	ACAGATTCTC	CTTCTAGTAT	GTATAGCGCA	AAAAGCTTCT	CGAGAGTACT	1140
TTAGAGACGG	CGCGGGGCC	ATCATTTTC	CACCCGGGTG	GGGTACC		1187

CCCTCACTAA AGGGAACAAA AGCTGGAGCT CGCGCGCCTG CAGGTCGACA CTAGTGGATC 60
GAAAGTTTCGT TACGCCAAGC TCGAAATTAA CTCTGGGCTG ACCCATAAAC ATTTGTCTGA 120

TCTAGGATAT	AGTTGCGTTT	CTTGCGGGCA	GCAATCTGGA	TGAGGCGGTT	GAGGCACTGG	180
GTGGCCTGCT	GGATCAGGAC	ATCCAGCGG	CCAGCATAGT	TCCGCTGCCG	GCCTAGGCCC	240
ATCACCCGCA	TCTTATCCAT	GATGGCATTG	GTACCCAGGA	TGTTGTACTT	CTTGGGAAGGG	300
TTGGAGGCTG	CATGTTTGAT	GGCCCATGTG	GTCTTGCCAG	CAGCAGGCAG	GCCCACCATC	360
ATCAGAACTC	CACATTCTGC	CTTGCTCTTT	GGTCCAACGG	TGCCCCGGAT	ACGCTCACTA	420
AGGGGAAGGT	GCTGGATGAA	GGTAAACCCC	GGGAGGACAG	AACAGTAGGG	CTCTGCTCTC	480
TGTCGGAAGT	TGAACTCCAC	TGCGCAATTC	TTCACCAGGA	CATGAGGATA	GAGGGCCTGA	540
CCCCCAAGG	CTTCCTTCTG	GATTGCGAAA	GCAATGCCCA	TCCACTTTCC	ATTCTTGGTG	600
AAAGACAGTT	CCACGTCAAT	TCCACATTCA	AAATCCGCAA	AGCAGCCAAT	CACCGGAGAG	660
CTCTGCGGTG	CTAGGAGAGC	GGCTGGGCCC	GCAGACTGGG	GGGAAAGCTC	CGCAGCCGCA	720
GTGGGCCCCA	GGATCAGGCC	CCGCGTGGCC	TGGAGAAGCC	CAGTCTGGGC	TGGAGCGGGA	780
GCTGGACAGT	GTGGCCTTGC	GTTCCGCCCC	GGGAGCGCTG	CGAGTGTGCG	GGCCTCGGGT	840
GGATTGTCTG	AGCACCAATA	CCTCACGGTT	GCCAACCTGG	GGTTTTAGCT	CCCTTGCTTT	900
TAATCCCCCTA	GGGGCGGGTG	GGGGCACGGG	AGGAAGGATG	GGCCAGCTGG	GTGCAATCCT	960
GCTGTAAGCT	AGCCATTCTT	TGATTCTTGA	GAATTAACCTA	AACGGTCGCG	CCGGAGGCGG	1020
CGGGGGCCGG	AGCGGAGCAG	CCGCGGCTGA	GGTTCGCGAG	TCGGCCGCTC	GGGGCTGCGC	1080
TCCCGCCGCG	GGACCCCGGC	CTCTGGCCGC	GCCGGCTCCG	GCCTCCGGGG	GGGCGGGGCG	1140
CGCCGGGACA	TGGTGCCAGT	CGCACCCCTT	CCCCGCCGCC	GCTGAGCTCG	CCGGCCGCGC	1200
CCGGGCTGGG	ACGTCCGAGC	GGGAAGATGT	TTTCCGCCCT	GAAGAAGCTG	GTGGGGTCGG	1260
ACCAGGCCCC	GGGCGGGGAC	AAGAACATCC	CCGCCGGGCT	GCAGTCCATG	AACCAGGCGT	1320
TGCAGAGGCG	CTTCGCCAAG	GGGGTGACGT	ACAACATGAA	GATAGTGATC	CGGGGAGACA	1380
GGAACACGGG	CAAGACAGCG	CTGTGGCACC	GCCTGCAGGG	CCGGCCGTTT	GTGGAGGAGT	1440
ACATCCCCAC	ACAGGAGATC	CAGGTACCCA	GCATCCACTG	GAGCTACAAG	ACCACGGATG	1500
ACATCGTGAA	GGTTGAAGTC	TGGGATGTAG	TAGACAAAGG	AAAATGCAAA	AAGCGAGGCG	1560
ACGGCTTAAA	GATGGAGAAC	GACCCCCAGG	AGNCGGAGTC	TGAAATGGCC	CTGGATGCTG	1620
AGTTGCTGAA	CGTGATACAG	AACTGCAACG	GGGTGGTCAT	GATGTTTGAC	ATTACCAAGC	1680
AGTGGACCTT	CAATTACATT	CTCCGGGAGC	TTCCAAAAGT	GCCCCACCAC	GTGCCAGTGT	1740
GCGTGCTGGG	GAATACCGG	GACATGGGCG	AGCACCGAGT	CATCCTGCCG	GACGACGTGC	1800
GTGACTTGAT	CGACAACCTG	GACAGACCTC	CAGGTTCTCT	CTACTTCCGC	TATGCTGAGT	1860
CTTCCATGAA	GAACAGCTTC	GGCCTAAAGT	ACCTTCATAA	GTTCTTCAAT	ATCCCATTYT	1920
TGCAGCTTCA	GAGGGAGACG	CTGTTGCGGC	AGCTGGAGAC	GAACCAGCTG	GACATGGACG	1980
CCACGCTGGA	GGAGCTGTCG	GTGCAGCAGG	AGACGGAGGA	CCAGAACTAC	GGCATCTTCC	2040
TGGAGATGAT	GGAGGCTCGC	AGCCGTGGCC	ATGCGTCCCC	ACTGGCGGCC	AACGGGCAGA	2100
GCCCATCCCC	GGGCTCCAG	TCACCACTCC	TGCCTGCACC	CGCTGTGTCC	ACGGGGAGCT	2160
CCAGCCCCGG	CACACCCGAC	CCCGCCCCAC	AGCTGCCCTT	CAATGCTGCC	CCACCATCCT	2220
CTGTGCCCGG	TGTACCAACC	TCAGAGGCCG	TGCCCCACCC	TGCGTGCCCC	TCAGCCCCCG	2280
CCCCACGGCG	CAGCATCATC	TCTAGGCTGT	TTGGGACGTC	ACCTGCCACC	GAGGCAGCCC	2340
CTCCACCTCC	AGAGCCAGTC	CCGGCCGCAC	AGGGCCGAGC	AACGGTCCAG	AGTGTGGAGG	2400
ACTTTGTTCC	TGACGACGCG	CTGGACCGCA	GCTTCCTGGA	AGACACAACC	CCCGCCAGGG	2460
ACGAGAAGAA	GGTGGGGGCC	AAGGCTGCCC	AGCAGGACAG	TGACAGTGAT	GGGGAGGGCC	2520
TGGGCGGGCA	CCCAGTGGTG	GCAGGGTTCC	AGGACGATGT	GGACCTCGAA	GACCAGCCAC	2580
GTGGGAGTCC	CCCCTGCTCT	GCAGGCCCCG	TCCCCAGTCA	AGACATCACT	CTTTCGAGTG	2640
AGGAGGAAGC	AGAACTGGCA	GCTCCACAAA	AAGGCCCTGC	CCCAGCTCCC	CAGCAGTGCT	2700
CAGAGCCAGA	GACCAAGTGG	TCCTCCATAC	CAGCTTCGAA	GCCACGGAGG	GGGACAGCTC	2760
CCACGAGGAC	CGCAGCACCC	CCCTGGCCAG	GCGGTGTCTC	TGTTGCGACA	GGTCCGGAGA	2820
AGCGCAGCAG	CACCAAGGCC	CCTGCTGAGA	TGGAGCCGGG	GAAGGGTGAG	CAGGCCTCCT	2880
CGTCGGAGAG	TGACCCCGAG	GGACCCATTG	CTGCACAAAT	GCTGTCCTTC	GTCATGGATG	2940
ACCCCCGACTT	TGAGAGCGAG	GGATCAGACA	CACAGCGCAG	GGCGGATGAC	TTTCCCGTGC	3000
GAGATGACCC	CTCCGATGTC	ACTGACGAGG	ATGAGGGCCC	TGCCGAGCCG	CCCCACCCCC	3060
CCAAGCTCCC	TCTCCCGGCC	TTGAGACTGA	AGAATGACTC	GGACCTCTTC	GGGCTGGGGC	3120
TGGAGGAGGC	CGGACCCGAG	GAGAGCAGTG	AGGAAGGTAA	GGAGGGCAAA	ACCCCTCTTA	3180
AGGAGAAGAA	AAAAAAACA	AAAAGCTTCT	CGAGAGTACT	TCTAGAGCGG	CCGCGGGCCC	3240
ATCGATTTTC	CACCGGGGTG	GGGTACCAGG	TAAGTGATACC	CAATTCGCCC	TATAGTGAGT	3300
CGTATT						3306

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TGCGGGGCCA GAGTGGGCTG

20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

GCAGTCCTGG CCTGCGGATG 20

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

GTCGACAGGA GAATTGGTTC 20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GCCTGGGTTC GGTGCGGGAC 20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

TGGTCGGGTG TTTGTGAGTG 20

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

CCTCTTCCGT CTCCTCAGTG 20

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

GGATTGCTAG TCTCACAGAC 20

[illegible]

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TTAAGGGTGG CTGAAGGGAC

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ACCTTCCCTC CCTGTCACAG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TGGTCGGGTG TTTGTGAGTG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ACACCATTCC AGAAATTCAG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

AAACTGCAGG TGGCTGAGTC

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GTCCTAATGT TTTCAGGGAG

20

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AAACCTATG GTTACAATTC

20

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TCCTAGACAT GGTCAAGTG

20

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GATATAATTA GTTCTCCATC

20

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ATGCCTGTTC CAGGCTGCAC

20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GGACGGCGAC CTCCACCCAC

20

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

GGGCTCCTCC GACGCCTGAG

20

(i) SEQUENCE CHARACTERISTICS:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AGTCTAGCCC TGGCCTTGAC

20

(i) SEQUENCE CHARACTERISTICS:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTCACTGGGG ACTCCGGCAG

20

(i) SEQUENCE CHARACTERISTICS:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CAGCTTTCCC TGGGCACATG

20

(i) SEQUENCE CHARACTERISTICS:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CACAGCTGTC TCAAGCCAG

20

(i) SEQUENCE CHARACTERISTICS:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ATCATATCCT CTTGCTGGTC

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GTTCCCAGAG CTTGTCTGTG

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GTGGGCGAGA CTCATAGTTG

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TAGCAGGGAG CCATGACCTG

20

(i) SEQUENCE CHARACTERISTICS:

- 1) SEQUENCE CHARACTERISTICS
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

CTTGGCGCCA GAAGCGAGAG

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CCTCTCTCTC TCTCTCTCTC

20

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

TCCCCGCTGA TTCGCCAAG

20

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CTTTTGAAT TCGGCACGAG

20

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CCCCTGGTCC GCACCAAGTTC

20

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GAGAAGGGTC GGGGCGGCAG

20

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

AAATCACATC GCGTCAACAC

20

(2) INFORMATION FOR SEQ ID NO:46:

0974394-4300
DOCTF:CTGCTAG

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TAAGAGAGTC ATAGTTACTC

20

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GCTCTAGAAG TACTCTCGAG

20

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ACTCTGGCCA TCAGGAGATC

20

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CAGGCGTTGT AGATGTTCTG

20

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGTGGCAGGC AGAAGTAATG

20

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

THE

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs

[illegible]

TGGGTGATGC CTTTGCTGAC

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

AAAACAAGAT CAAGGTGATG

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTGCCACAT TGCTATGGTG

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

GACCAAGATC AGAAGTAGAG

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CCCCTGGGCC AATGATGTTG

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

TCTTCCCACC ATAGCAATG

19

[illegible]

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TGGTCTTGGT GACCAATGTG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ACACCTCGGT GACCCCTGTG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TCTCCAAGTT CGGCACAGTG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ACATGGGGCTG CACTCACGAC

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GATCCTCTGA ACCTGCAGAG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

00712043 44300
00E11:ET021260

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:
GGAAATGAGG TGGGCGATC 20
(2) INFORMATION FOR SEQ ID NO:69:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:
CTTGCCTTG GACAAGGATG 20
(2) INFORMATION FOR SEQ ID NO:70:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:
GCACCTGCCA TTGGGGGTAG 20
(2) INFORMATION FOR SEQ ID NO:71:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:
GGTGGAAGCC ATTGACGGTG 20
(2) INFORMATION FOR SEQ ID NO:72:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:
TGCCTCTCTC GTCGCTGCTG 20
(2) INFORMATION FOR SEQ ID NO:73:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:
GCGGAAACTC TGTGGTGCTG 20
(2) INFORMATION FOR SEQ ID NO:74:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

AGGATTGCCT TCCTCTACTG

20

(2) INFORMATION FOR SEQ ID NO:75:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

TGTCTGTTTC ACCAGGGCAG

20

(2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

CCAGTGCCTC TATGCATGTC

20

(2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

AGGAAGCCCA CGCACACCAC

20

(2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CCCTTTGTTT CCTGATCTTC

20

(2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

CGCTCGGGAT CCAGGTCATC

20

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

TCGAGGTTCA GAGCGTAGTG

20

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

TCTTGATCT CTGGCACCTC

20

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

CCATCAGAGT GAAGGAGGAG

20

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

CCATCTTCCA CTGGTCAGAG

20.

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

CTCCTTCTCT TGGATCTCTG

20

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid

[illegible]

TTACTTCAGC ACTGTTAGTC 20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

AGGGAGGTAG CTCAAAGCTC 20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TGGGTCCACA GTTCGCACAG 20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

CAACTCTGTG ATGGCTCCAG 20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

AGCAGGGTTC TGTTCAAGAC 20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

CCATTGGGTG CTAGTCTCTC 20

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

CAGCCATGCT GTCCAGCAG

20

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

CTGGACCTGA GGTAGCGCTG

20

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

ATAACCACCC TGAGGCACTG

20

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

CCTGCAGGTC GACACTAGTG

20

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

AATTGGAATG AGGAGGACTG

20

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

Figure 1 displays 15 chemical structures, labeled (a) through (o), which are substituted benzene derivatives. The structures are arranged vertically and include various functional groups and substituents, such as alcohols, aldehydes, ketones, and esters, along with different ring substituents like methyl, ethyl, and propyl groups.

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(i) SEQUENCE CHARACTERISTICS:

[illegible]

TTGTACACCA GCTGCAGGTC

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GGGTGTGGTG CAGATGAGTC

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ATCACACTCT TATAGCTCAG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GTGGGAAGCT TTCCTCAGAC

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TGATGAACAT GGGCCTGGAG

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

— 73 —

20

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

20

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

20

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

20

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

20

(i) SEQUENCE CHARACTERISTICS:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

20

(i) SEQUENCE CHARACTERISTICS:

- 74 —

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

GCACGAGTGG CACAAGCGTG

20

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

GCAAGCGTGT GGTGTCAGTG

20

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

TGTTTGAACA GGCTCTGGAC

20

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

CGGCATGGCA ATGAGGACAC

20

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

AGGACGAGAT GGACCTCCAG

20

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

CCCTCTGTCC TCTAGCCAC

20

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

TCTTGAGGGG ACTGACTCTG

20

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

TGAGTGAGGA GGCAGATGTC

20

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

TGGCTTTGAA GAAAGAGCTG

20

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

GCAAAAGACC AGGCTGACTG

20

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

TGCAGCTCCT TGGTCTTCTC

20

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

[illegible]

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20_base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20

(A) LENGTH: 20 base pairs

TGGTCTCTGG CTCTGAGCAC

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GCCTGGAGAA GCCCAGTCTG

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

CACACTCTGG ACCGTTGCTG

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

AAAGCTCCGC AGCCGCAGTG

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TCTTCCAGGA AGCTGCGGTC

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

GATGGTGGGG CAGCATTGAG

20

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GTCACCAGTG GTGCCTGCAG

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ACCTCACGGT TGCCAACCTG

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

CGCAACAGCG TCTCCCTCTG

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

AGTACCTTCA TAAGTTCTTC

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

TCCCAGACTT CAACCTTCAC

20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

[illegible]

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GCTGAGCACC TTTACCTCAC

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GACGTCCGTC CGGGAAGATG

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

ACACAGGAGA TGCAGGTCAC

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GAGTCTTCCA TGAAGAACAG

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

GCAGTGAGGA AGGTAAGGAG

20

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4047 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 378...1799
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

```
GGATCCAAAG GACGCCCCCG CCGACAGGAG AATTGGTTCC CGGGCCCGCG GCGATGCCCC 60
CCCGGTAGCT CGGGCCCGTG GTCGGGTGTT TGTGAGTGTT TCTATGTGGG AGAAGGAGGA 120
GGAGGAGGAA GAAGAAGCAA CGATTGTCT TCTCGGCTGG TCTCCCCCGG GCTCTACATG 180
TTCCCCGCAC TGAGGAGACG GAAGAGGAGC CGTAGCCGCC CCCCCTCCCG GCCCGGATTA 240
TAGTCTCTCG CCACAGCGGC CTCGGCCTCC CTTGGATTG AGACGCCGAT TCGCCCAAGT 300
TTTGGGAAAT GGAAGTAAT GACAGCTGGC ACCTGAAC TAAGTACTTTA TAGGCAACAC 360
CATTCCAGAA ATTCAGG ATG AAT GGG GAT ATG CCC CAT GTC CCC ATT ACT 410
          Met Asn Gly Asp Met Pro His Val Pro Ile Thr
              1              5              10

ACT CTT GCG GGG ATT GCT AGT CTC ACA GAC CTC CTG AAC CAG CTG CCT 458
Thr Leu Ala Gly Ile Ala Ser Leu Thr Asp Leu Leu Asn Gln Leu Pro
              15              20              25

CTT CCA TCT CCT TTA CCT GCT ACA ACT ACA AAG AGC CTT CTC TTT AAT 506
Leu Pro Ser Pro Leu Pro Ala Thr Thr Thr Lys Ser Leu Leu Phe Asn
              30              35              40

GCA CGA ATA GCA GAA GAG GTG AAC TGC CTT TTG GCT TGT AGG GAT GAC 554
Ala Arg Ile Ala Glu Glu Val Asn Cys Leu Leu Ala Cys Arg Asp Asp
              45              50              55

AAT TTG GTT TCA CAG CTT GTC CAT AGC CTC AAC CAG GTA TCA ACA GAT 602
Asn Leu Val Ser Gln Leu Val His Ser Leu Asn Gln Val Ser Thr Asp
              60              65              70              75

CAC ATA GAG TTG AAA GAT AAC CTT GGC AGT GAT GAC CCA GAA GGT GAC 650
His Ile Glu Leu Lys Asp Asn Leu Gly Ser Asp Asp Pro Glu Gly Asp
              80              85              90

ATA CCA GTC TTG TTG CAG GCC GTC CTG GCA AGG AGT CCT AAT GTT TTC 698
Ile Pro Val Leu Leu Gln Ala Val Leu Ala Arg Ser Pro Asn Val Phe
              95              100              105

AGG GAG AAA AGC ATG CAG AAC AGA TAT GTA CAA AGT GGA ATG ATG ATG 746
Arg Glu Lys Ser Met Gln Asn Arg Tyr Val Gln Ser Gly Met Met Met
              110              115              120

TCT CAG TAT AAA CTT TCT CAG AAT TCC ATG CAC AGT AGT CCT GCA TCT 794
Ser Gln Tyr Lys Leu Ser Gln Asn Ser Met His Ser Ser Pro Ala Ser
              125              130              135

TCC AAT TAT CAA CAA ACC ACT ATC TCA CAT AGC CCC TCC AGC CGG TTT 842
Ser Asn Tyr Gln Gln Thr Thr Ile Ser His Ser Pro Ser Ser Arg Phe
              140              145              150              155

GTG CCA CCA CAG ACA AGC TCT GGG AAC AGA TTT ATG CCA CAG CAA AAT 890
Val Pro Pro Gln Thr Ser Ser Gly Asn Arg Phe Met Pro Gln Gln Asn
              160              165              170

AGC CCA GTG CCT AGT CCA TAC GCC CCA CAA AGC CCT GCA GGA TAC ATG 938
Ser Pro Val Pro Ser Pro Tyr Ala Pro Gln Ser Pro Ala Gly Tyr Met
              175              180              185

CCA TAT TCC CAT CCT TCA AGT TAC ACA ACA CAT CCA CAG ATG CAA CAA 986
Pro Tyr Ser His Pro Ser Ser Tyr Thr Thr His Pro Gln Met Gln Gln
              190              195              200
```

GCA Ala	TCG Ser	GTA Val	TCA Ser	AGT Ser	CCC Pro	ATT Ile	GTT Val	GCA Ala	GGT Gly	GGT Gly	TTG Leu	AGA Arg	AAC Asn	ATA Ile	CAT His	1034																			
GAT Asp	AAT Asn	AAA Lys	GTT Val	TCT Ser	GGT Gly	CCG Pro	TTG Leu	TCT Ser	GGC Gly	AAT Asn	TCA Ser	GCT Ala	AAT Asn	CAT His	CAT His	1082																			
GCT Ala	GAT Asp	AAT Asn	CCT Pro	AGA Arg	CAT His	GGT Gly	TCA Ser	AGT Ser	GAG Glu	GAC Asp	TAC Tyr	CTA Leu	CAC His	ATG Met	GTG Val	1130																			
CAC His	AGG Arg	CTA Leu	AGT Ser	AGT Ser	GAC Asp	GAT Asp	GGA Gly	GAT Asp	TCT Ser	TCA Ser	ACA Thr	ATG Met	AGG Arg	AAT Asn	GCT Ala	1178																			
GCA Ala	TCT Ser	TTT Phe	CCC Pro	TTG Leu	AGA Arg	TCT Ser	CCA Pro	CAG Gln	CCA Pro	GTA Val	TGC Cys	TCC Ser	CCT Pro	GCT Ala	GGA Gly	1226																			
AGT Ser	GAA Glu	GGA Gly	ACT Thr	CCT Pro	AAA Lys	GGC Gly	TCA Ser	AGA Arg	CCA Pro	CCT Pro	TTA Leu	ATC Ile	CTA Leu	CAA Gln	TCT Ser	1274																			
CAG Gln	TCT Ser	CTA Leu	CCT Pro	TGT Cys	TCA Ser	TCA Ser	CCT Pro	CGA Arg	GAT Asp	GTT Val	CCA Pro	CCA Pro	GAT Asp	ATC Ile	TTG Leu	1322																			
CTA Leu	GAT Asp	TCT Ser	CCA Pro	GAA Glu	AGA Arg	AAA Lys	CAA Gln	AAG Lys	AAG Lys	CAG Gln	AAG Lys	AAA Lys	ATG Met	AAA Lys	TTA Leu	1370																			
GGC Gly	AAG Lys	GAT Asp	GAA Glu	AAA Lys	GAG Glu	CAG Gln	AGT Ser	GAG Glu	AAA Lys	GCG Ala	GCA Ala	ATG Met	TAT Tyr	GAT Asp	ATA Ile	1418																			
ATT Ile	AGT Ser	TCT Ser	CCA Pro	TCC Ser	AAG Lys	GAC Asp	TCT Ser	ACT Thr	AAA Lys	CTT Leu	ACA Thr	TTA Leu	AGA Arg	CTT Leu	TCT Ser	1466																			
CGT Arg	GTA Val	AGG Arg	TCT Ser	TCA Ser	GAC Asp	ATG Met	GAC Asp	CAG Gln	CAA Gln	GAG Glu	GAT Asp	ATG Met	ATT Ile	TCT Ser	GGT Gly	1514																			
GTG Val	GAA Glu	AAT Asn	AGC Ser	AAT Asn	GTT Val	TCA Ser	GAA Glu	AAT Asn	GAT Asp	ATT Ile	CCT Pro	TTT Phe	AAT Asn	GTG Val	CAG Gln	1562																			
TAC Tyr	CCA Pro	GGA Gly	CAG Gln	ACT Thr	TCA Ser	AAA Lys	ACA Thr	CCC Pro	ATT Ile	ACT Thr	CCA Pro	CAA Gln	GAT Asp	ATA Ile	AAC Asn	1610																			
CGC Arg	CCA Pro	CTA Leu	AAT Asn	GCT Ala	GCT Ala	CAA Gln	TGT Cys	TTG Leu	TCG Ser	CAG Gln	CAA Gln	GAA Glu	CAA Gln	ACA Thr	GCA Ala	1658																			
TTC Phe	CTT Leu	CCA Pro	GCA Ala	AAT Asn	CAA Gln	GTG Val	CCT Pro	GTT Val	TTA Leu	CAA Gln	CAG Gln	AAC Asn	ACT Thr	TCA Ser	GTT Val	1706																			
GCT Ala	GCA Ala	AAA Lys	CAA Gln	CCC Pro	CAG Gln	ACC Thr	AAT Asn	AGT Ser	CAC His	AAA Lys	ACC Thr	TTG Leu	GTG Val	CAG Gln	CCT Pro	1754																			
GGA Gly	ACA Thr	GGC Gly	ATA Ile	GAG Glu	GTG Val	TCA Ser	GCA Ala	GAG Glu	CTG Leu	CCC Pro	AAG Lys	GAC Asp	AAG Lys	ACC Thr	TAAGA Thr	1804																			
TCCAGCAGGG	AAC	TATGTAG	TCACCCCGAG	AGGCCAGCT	CTCTCCGTGA	GCTCTGGGCC	1864	TAGGGTGGGG	GTGGTTGTTG	GTTCTGCGCG	CACTGTTCCC	CCTACATGAT	GGGTCCATCC	1924	CAGTTGGCTT	CTCTCACTCG	CTTCCTCTG	TGGAGAAGCC	TGTCACAGTG	TCACCTGCCTC	1984	CAGGAAGCTG	TCTCTGATTT	CTCCAGTTGA	ACAGTGAGAT	TTGCCACACC	TCACATGCAT	2044	CGCTCTTGTC	CCTGGAATTG	TAACCATAGG	TTTTCTGTG	TCCTGGAGGA	CAAGGATGAG	2104

GGCTTTCCAC TTGAGTCTCC CTGGTGGAGC CCAGCTCCTG ACATACCTGG TAAAAGTTCT 2164
 CAAGAGAAGA ACATGGAGGA GGAATGTGGA TAACAACCTT GGCTGCCTGT GTGTTCCAAG 2224
 CTAGGAAGAT GTAATGTCCC CACAAACGGG GTAAATGGCT TGCCTGCGTC ACAGCTGTCT 2284
 CAAGCCCAGG CCCTGGGCGC CAGCCCAAGC CCAAGGACTA GGTCCAGAGC CACACAGCGC 2344
 CAGGCCACAT CCGCCTCACC TGGGACCCTT TGTGGGGTAC AGTCTCCGGC CCCACCCAGA 2404
 CCTCCTGAAG GAGAGACCCC ATGGCAAGGA CTCAGCCACC TGCAGTTTCA TAAGCCCCCA 2464
 GTGGGTTCCT AGGCATGAAG ACCACCGGTT AGAGGCTGAA CTGGCAGGAA CCTGTCTCCA 2524
 GCCCCTTCTC ACCCCAGCCG GGCCCTGCCT CAGAGGCAGC ACCCAGGACG TGGCCATGAC 2584
 CCGTGGGACTC CACTCAATCC CTCTTCTCCA GGAGCCATGC AAAGTGTCAG CCAGCCAGGC 2644
 CCCTGGAAGG CAGTCATCAC CTCTTAAGGC ATTGTGGGTG TCGGTCCTGC AACTGCCAGG 2704
 TGCAGCACAC GACCCGTGTC CGGTGTTTGA TAGCAGGGAG CCATGACCTG GCAACGATTC 2764
 CACGCTCAAA GGGGCACCCG GGGGGCCCTG GGTGCGGGCG GATCAGCTTT CCCTGGGCAC 2824
 ATCTGCCTCA TTCCAGATCT CCAGGGCTCA TGTCTGTGAC AGGGAGGGAA GGCTCTGCCC 2884
 TGGCCTTCCG TCAGCTCTGC CAGTGCAGGC TGGGCAGCCT GGGCTTTAGA GCTGGCTTCT 2944
 GCCCACACTT TCTCCGTGAA AGGAAAACAA CTATGAGTCT GCCAAACGCA TCTCAGATGC 3004
 GTTTTAAAAA ATTCTGGTCC CCGCTCTCTG TCCCATCATC CGCCTCGGGG ACTTCCTCTC 3064
 TCCGTGGTTC TCACCCATA CTCTGTCACT GCCACATTTT CACCTGGGCC TGGCCTTTGT 3124
 CTCACCTGA AACTCCTGAA AATCTTGAAA TGGATTCTA GGTCACTGGG GACTCCGGCA 3184
 GCACATTCGG CTTCAGAATA AAGGGCGCCC GCGGTCCCCC AGCACCTCCC CAAGCCACAC 3244
 CCCTAGCTTC CCTCCCTATC CTGTCAGCCT GAGGGTCCCT TCAGCCACCC TTAAGTCCCC 3304
 ACCTGGGCTC CTGCCCCGCC CTGGGCTAGC AGCGCCTTCT CCACCGGGGC CCCCTCTGCT 3364
 CACAGAGCCC CCTCACCTCC CTGGGGATGA GGGGCCAGGC CATGACCTG AAAGTCTAGC 3424
 CCTGGCCTTG ACCTCCAGG AGCGCCCTCC CCGCCCTCTC CCGGCCCCGG CCCCGTCTC 3484
 TGCTGCTGGC CTCTGGGTG TCCCCGCGAG ACTGAGCTGC GCTTGGGGGT CCTGGCGGCC 3544
 TGGGCCGTCC CGCACCGAAC CCAGGCGGTC GGAGCCCGC GGGAAAGGCG GAGGTCCTTC 3604
 TGGGGCTCC TCCGACGCT GAGGGCGCTG CTCCCCCGG GCCGCCCCGG GTTCTGCGG 3664
 AGCCGGGGCC TCCGCTCTG GGTGACCCGG TGAGACCCCC GGGGAGGCG CTGGGGAGGC 3724
 GCGGGCTCTG CTCCCCGGT CCAAACGCAC TGGCTGCCCC TCAGGAGGGA CGGCGACCTC 3784
 CACCCACGGC GCTGGCGCCC GCACGGCCGC TCCTCCCGCT CCCGAGCCT GGACGCTCC 3844
 CGAGGCCGCC CCGCCGGGCC CCACGCGCGG CCCCATCCGC AGGCCAGGAC TGCCTTCCCG 3904
 GAGCTGGCGG CCCCCAGCCT GGAGGAGCCG GCCCCAGAGC CCCTCCAGC CCTCCCCAGC 3964
 CCACTCTGGC CCGCAGGCC CCGCCTGGT CGAGTGCGGG TCTCTGGCCC GGGCCTTTCC 4024
 CGGGGAAGGA AAGCAAAAG CTT 4047

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 474 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

Met Asn Gly Asp Met Pro His Val Pro Ile Thr Thr Leu Ala Gly Ile
 1 5 10 15
 Ala Ser Leu Thr Asp Leu Leu Asn Gln Leu Pro Leu Pro Ser Pro Leu
 20 25 30
 Pro Ala Thr Thr Thr Lys Ser Leu Leu Phe Asn Ala Arg Ile Ala Glu
 35 40 45
 Glu Val Asn Cys Leu Leu Ala Cys Arg Asp Asp Asn Leu Val Ser Gln
 50 55 60
 Leu Val His Ser Leu Asn Gln Val Ser Thr Asp His Ile Glu Leu Lys
 65 70 75 80
 Asp Asn Leu Gly Ser Asp Asp Pro Glu Gly Asp Ile Pro Val Leu Leu
 85 90 95
 Gln Ala Val Leu Ala Arg Ser Pro Asn Val Phe Arg Glu Lys Ser Met
 100 105 110
 Gln Asn Arg Tyr Val Gln Ser Gly Met Met Ser Gln Tyr Lys Leu
 115 120 125
 Ser Gln Asn Ser Met His Ser Ser Pro Ala Ser Ser Asn Tyr Gln Gln
 130 135 140
 Thr Thr Ile Ser His Ser Pro Ser Ser Arg Phe Val Pro Pro Gln Thr
 145 150 155 160
 Ser Ser Gly Asn Arg Phe Met Pro Gln Gln Asn Ser Pro Val Pro Ser
 165 170 175
 Pro Tyr Ala Pro Gln Ser Pro Ala Gly Tyr Met Pro Tyr Ser His Pro
 180 185 190
 Ser Ser Tyr Thr Thr His Pro Gln Met Gln Gln Ala Ser Val Ser Ser
 195 200 205

[illegible]

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2998 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 26...799
(C) OTHER INFORMATION:

AAGCTTTT	TG	AATTCGGCAC	GAGAT	GCT	ACA	CAG	GCT	ATA	TTT	GAA	ATA	CTG	52			
				Ala	Thr	Gln	Ala	Ile	Phe	Glu	Ile	Leu				
				1				5								
GAG	AAA	TCC	TGG	TTG	CCC	CAG	AAT	TGT	ACA	CTG	GTT	GAT	ATG	AAG	ATT	100
Glu	Lys	Ser	Trp	Leu	Pro	Gln	Asn	Cys	Thr	Leu	Val	Asp	Met	Lys	Ile	
10					15					20					25	
GAA	TTT	GGT	GTT	GAT	GTA	ACC	ACC	AAA	GAA	ATT	GTT	CTT	GCT	GAT	GTT	148
Glu	Phe	Gly	Val	Asp	Val	Thr	Thr	Lys	Glu	Ile	Val	Leu	Ala	Asp	Val	
				30					35					40		
ATT	GAC	AAT	GAT	TCC	TGG	AGA	CTC	TGG	CCA	TCA	GGA	GAT	CGA	AGC	CAA	196
Ile	Asp	Asn	Asp	Ser	Trp	Arg	Leu	Trp	Pro	Ser	Gly	Asp	Arg	Ser	Gln	
			45					50					55			
CAG	AAA	GAC	AAA	CAG	TCT	TAT	CGG	GAC	CTC	AAA	GAA	GTA	ACT	CCT	GAA	244
Gln	Lys	Asp	Lys	Gln	Ser	Tyr	Arg	Asp	Leu	Lys	Glu	Val	Thr	Pro	Glu	
		60					65					70				
GGG	CTC	CAA	ATG	GTA	AAG	AAA	AAC	TTT	GAG	TGG	GTT	GCA	GAG	AGA	GTA	292

Gly	Leu	Gln	Met	Val	Lys	Lys	Asn	Phe	Glu	Trp	Val	Ala	Glu	Arg	Val	
75						80					85					
GAG	TTG	CTT	TTG	AAA	TCA	GAA	AGT	CAG	TGC	AGG	GTT	GTA	GTG	TTG	ATG	340
Glu	Leu	Leu	Leu	Lys	Ser	Glu	Ser	Gln	Cys	Arg	Val	Val	Val	Leu	Met	
90					95					100					105	
GGC	TCT	ACT	TCT	GAT	CTT	GGT	CAC	TGT	GAA	AAA	ATC	AAG	AAG	GCC	TGT	388
Gly	Ser	Thr	Ser	Asp	Leu	Gly	His	Cys	Glu	Lys	Ile	Lys	Lys	Ala	Cys	
				110					115					120		
GGA	AAT	TTT	GGC	ATT	CCA	TGT	GAA	CTT	CGA	GTA	ACA	TCT	GCG	CAT	AAA	436
Gly	Asn	Phe	Gly	Ile	Pro	Cys	Glu	Leu	Arg	Val	Thr	Ser	Ala	His	Lys	
			125					130					135			
GGA	CCA	GAT	GAA	ACT	CTG	AGG	ATT	AAA	GCT	GAG	TAT	GAA	GGG	GAT	GGC	484
Gly	Pro	Asp	Glu	Thr	Leu	Arg	Ile	Lys	Ala	Glu	Tyr	Glu	Gly	Asp	Gly	
	140					145						150				
ATT	CCT	ACT	GTA	TTT	GTG	GCA	GTG	GCA	GGC	AGA	AGT	AAT	GGT	TTG	GGA	532
Ile	Pro	Thr	Val	Phe	Val	Ala	Val	Ala	Gly	Arg	Ser	Asn	Gly	Leu	Gly	
	155					160					165					
CCA	GTG	ATG	TCT	GGG	AAC	ACT	GCA	TAT	CCA	GTT	ATC	AGC	TGT	CCT	CCC	580
Pro	Val	Met	Ser	Gly	Asn	Thr	Ala	Tyr	Pro	Val	Ile	Ser	Cys	Pro	Pro	
170				175						180				185		
CTC	ACA	CCA	TAC	TGG	GGA	GTT	CAG	GAT	GTG	TGG	TCT	TCT	CTT	CGA	CTA	628
Leu	Thr	Pro	Asp	Trp	Gly	Val	Gln	Asp	Val	Trp	Ser	Ser	Leu	Arg	Leu	
				190					195					200		
CCC	AGT	GGT	CTT	GGC	TGT	TCA	ACC	GTA	CTT	TCT	CCA	GAA	GGA	TCA	GCT	676
Pro	Ser	Gly	Leu	Gly	Cys	Ser	Thr	Val	Leu	Ser	Pro	Glu	Gly	Ser	Ala	
			205					210					215			
CAA	TTT	GCT	GCT	CAG	ATA	TTT	GGG	TTA	AGC	AAC	CAT	TTG	GTA	TGG	AGC	724
Gln	Phe	Ala	Ala	Gln	Ile	Phe	Gly	Leu	Ser	Asn	His	Leu	Val	Trp	Ser	
	220					225					230					
AAA	CTG	CGA	GCA	AGC	ATT	TTG	AAC	ACA	TGG	ATT	TCC	TTG	AAG	CAG	GCT	772
Lys	Leu	Arg	Ala	Ser	Ile	Leu	Asn	Thr	Trp	Ile	Ser	Leu	Lys	Gln	Ala	
	235					240					245					
GAC	AAG	AAA	ATC	AGA	GAA	TGT	AAT	TTA	TAAGAAAGAA	TGCCATTGAA	TTTTTTA					826
Asp	Lys	Lys	Ile	Arg	Glu	Cys	Asn	Leu								
250					255											
GGGGAAAAAC	TACAAATTTT	TAATTTAGCT	GAAGGAAAAAT	CAAGCAAGAT	GAAGAGGTAA											886
TTTTAAATTA	GAGAACACAA	ATAAAATGTA	TTAGTGAATA	AATGGTGAGG	GTAGGCCAT											946
TCAGATGCAA	GGCCAGCAAT	GGGGCTCCCC	ATTATCCCCA	CCCTTTTGGT	CCAGTCCCC											1006
TTCTCTGCAA	TGGGCACGCA	TAGAGGAGAG	ACAAAGGGTA	TTAGACGCAA	CATCATTGGC											1066
CCAGGGGAGT	CCGAGAAGAG	CTGCCATTGG	CTGACAGGGC	ATTTTCAGGC	TCTGTCAATTG											1126
GTCAGGGAGC	ACACCCCAGC	CTGAAGAGTG	ATGCCATTGG	CCAGGGAGTG	GTTTTGTCAT											1186
AGGCGTTGGC	TGTGAAGTGG	AAGGAAAAGA	TCTGGGAATG	AAGCCCTGTG	GCCAGGAAGA											1246
TAGACAGGGC	AGCAACTTCT	GGGCCTCCAG	GCCCTCTTCC	CACCATAGCA	ATGTGGGCAA											1306
AACTTGGTGC	AGGCCCCAGC	CAGAAAAAAG	AGCCCAAGCC	AGAGGGGCAAG	TGACAAAGGA											1366
TGTACCATGT	CCAATCTCCC	ACACCCTGGG	GCTGCCCTTC	CCAATGTCTT	TCTTGATAGC											1426
CAAGTTGGGC	TGGGAGCAGC	TCACTGCTCC	TCTAGCCAGG	AGGGTTTCTC	AGCTCCTGGA											1486
GGCCGACAGT	TGATGTTGAA	CTGCTGCAGG	GTCTGCTCCA	GCTGTTTCTG	GTTCCCCAGCA											1546
AAGTAGGCGG	ACACAGCATT	GTGGAAGAGC	AGCAGCTGCT	TGTGCATCAC	CTTGATCTTG											1606
TTTTTCTTCCA	GGAACTTGAG	CTTGATGGCC	ACATCTCCCC	GAGCCTTCTC	ATACTTGTCC											1666
GGATGGGCCT	GGAAAGTGGC	CTGGGCCTCT	TCAAGTCGAC	CACGCTGTCC	TGCATCCCGG											1726
CCGCCTAGAC	TGACGCTCTC	TAAGTCTGTG	CGGTAGGCAT	CATATTCAGC	CCTGGCAGCC											1786
TCATACTGTT	TCACAGTCTA	GAGCGTGTCT	TCCATTGGCT	TGGTGACCAA	TGTTGTTGATG											1846
CTAGAGACAA	AGAAGTTTCA	GGCTCTTAGC	AGCGTITCCC	CATCTTGTCA	TAGTAGTTTC											1906
TGTGTCCTCTG	CATTGTAGCC	AAATTCTCTC	TGAAGCTCTG	GGGACTTCTG	GCTGAGGTCA											1966
GCAAAGGCAT	CACCCAGTGC	ATGCTGGGTC	TGCAGCAGGC	TGTAGAGGTG	GGCTGTCACT											2026
GCCCGGCCCA	GCTGCAGAGC	ACTCTCATAC	TTGCGCTTCG	TCTCAGCGAC	CAACTCAATC											2086
TGCAGCTCTA	GCTCCAGGAT	TCCGCGCGCT	CCACTCCGTC	GGCGAGGGT	GCTGCTGTGTG											2146
TGCCATTGAC	GGCATTTGCC	CAGATATAGC	CGTTGGTACA	AAGCGGGGAT	CTGACGAGCT											2206
TTTCTCTACT	TGTGTCACTA	ACGGACCGTT	TATCATGAGC	AGCAACTCGG	CTTCTGCAGC											2266
AAACGGAAAT	GACAGCAAGA	AGTTCAAAGG	TGACAGCCGA	AGTGCAGCGG	TCCCTCTAG											2326
AGTGATCCAC	ATCCGGAAGT	TCCCATCCGA	CGTCACGGAG	GGGGAAGTCA	TCTCCCTGGG											2386

GCTGCCCTTT GGAAGGTCA CCAACCTCCT GATGCTGAAG GGGAAAAACC AGGCCTTCAT 2446
CGAGATGAAC ACGGAGGAGG CTGCCAATAC CATGGTGAAC TACTACACCT CGGTGACCCC 2506
TGTGCTGCGC GGCCAGCCCA TCTACATCCA GTTCTCCAAC CACAAGGAGC TGAAGACCGA 2566
CAGCTCTCCC AACCAGGCGC GGGCCAGGC GGCCCTGCAG GCGGTGAACT CGGTCCAGTC 2626
GGGGAACCTG GCCTTGGCTG CCTCGGCGGC GGCCCTGGAT GCAGGGATGG CGATGGCCGG 2686
GCAGAGCCCC GTGCTCAGGA TCATCGTGA GAACCTCTTC TACCCTGTGA CCCTGGATGT 2746
GCTGCACCAG ATTTTCTCCA AGTTCGGCAC AGTGTGAAG ATCATCACCT TCACCAAGAA 2806
CAACCAGTTC CAGGCCCTGC TGCAGTATGC GGACCCCGTG AGCGCCAGC ACGCCAAGCT 2866
GTCGCTGGAC GGGCAGAACA TCTACAACGC CTGCTGCACG CTGCGCATCG ACTTTTCAA 2926
CTCACCAGC CTCAACGTCA AGTACAACAA TGACAAGAGC CGTGACTACC TCGTGCCGAA 2986
TCTTTGGAT CC 2998

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

Ala Thr Gln Ala Ile Phe Glu Ile Leu Glu Lys Ser Trp Leu Pro Gln
1 5 10 15
Asn Cys Thr Leu Val Asp Met Lys Ile Glu Phe Gly Val Asp Val Thr
20 25 30
Thr Lys Glu Ile Val Leu Ala Asp Val Ile Asp Asn Asp Ser Trp Arg
35 40 45
Leu Trp Pro Ser Gly Asp Arg Ser Gln Gln Lys Asp Lys Gln Ser Tyr
50 55 60
Arg Asp Leu Lys Glu Val Thr Pro Glu Gly Leu Gln Met Val Lys Lys
65 70 75 80
Asn Phe Glu Trp Val Ala Glu Arg Val Glu Leu Leu Lys Ser Glu
85 90 95
Ser Gln Cys Arg Val Val Val Leu Met Gly Ser Thr Ser Asp Leu Gly
100 105 110
His Cys Glu Lys Ile Lys Lys Ala Cys Gly Asn Phe Gly Ile Pro Cys
115 120 125
Glu Leu Arg Val Thr Ser Ala His Lys Gly Pro Asp Glu Thr Leu Arg
130 135 140
Ile Lys Ala Glu Tyr Glu Gly Asp Gly Ile Pro Thr Val Phe Val Ala
145 150 155 160
Val Ala Gly Arg Ser Asn Gly Leu Gly Pro Val Met Ser Gly Asn Thr
165 170 175
Ala Tyr Pro Val Ile Ser Cys Pro Pro Leu Thr Pro Asp Trp Gly Val
180 185 190
Gln Asp Val Trp Ser Ser Leu Arg Leu Pro Ser Gly Leu Gly Cys Ser
195 200 205
Thr Val Leu Ser Pro Glu Gly Ser Ala Gln Phe Ala Ala Gln Ile Phe
210 215 220
Gly Leu Ser Asn His Leu Val Trp Ser Lys Leu Arg Ala Ser Ile Leu
225 230 235 240
Asn Thr Trp Ile Ser Leu Lys Gln Ala Asp Lys Lys Ile Arg Glu Cys
245 250 255
Asn Leu

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1038 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

Ile Gln Arg Phe Gly Thr Ser Gly His Ile Met Asn Leu Gln Ala Gln
1 5 10 15

Pro Lys Ala Gln Asn Lys Arg Lys Arg Cys Leu Phe Gly Gly Gln Glu
 20 25 30
 Pro Ala Pro Lys Glu Gln Pro Pro Pro Leu Gln Pro Pro Gln Gln Ser
 35 40 45
 Ile Arg Val Lys Glu Glu Gln Tyr Leu Gly His Glu Gly Pro Gly Gly
 50 55 60
 Ala Val Ser Thr Ser Gln Pro Val Glu Leu Pro Pro Ser Ser Leu
 65 70 75 80
 Ala Leu Leu Asn Ser Val Val Tyr Gly Pro Glu Arg Thr Ser Ala Ala
 85 90 95
 Met Leu Ser Gln Gln Val Ala Ser Val Lys Trp Pro Asn Ser Val Met
 100 105 110
 Ala Pro Gly Arg Gly Pro Glu Arg Gly Gly Gly Gly Val Ser Asp
 115 120 125
 Ser Ser Trp Gln Gln Gln Pro Gly Gln Pro Pro Pro His Ser Thr Trp
 130 135 140
 Asn Cys His Ser Leu Ser Leu Tyr Ser Ala Thr Lys Gly Ser Pro His
 145 150 155 160
 Pro Gly Val Gly Val Pro Thr Tyr Tyr Asn His Pro Glu Ala Leu Lys
 165 170 175
 Arg Glu Lys Ala Gly Gly Pro Gln Leu Asp Arg Tyr Val Arg Pro Met
 180 185 190
 Met Pro Gln Lys Val Gln Leu Glu Val Gly Arg Pro Gln Ala Pro Leu
 195 200 205
 Asn Ser Phe His Ala Ala Lys Lys Pro Pro Asn Gln Ser Leu Pro Leu
 210 215 220
 Gln Pro Phe Gln Leu Ala Phe Gly His Gln Val Asn Arg Gln Val Phe
 225 230 235 240
 Arg Gln Gly Pro Pro Pro Pro Asn Pro Val Ala Ala Phe Pro Pro Gln
 245 250 255
 Lys Gln Gln Gln Gln Gln Gln Pro Gln Gln Gln Gln Gln Gln
 260 265 270
 Ala Ala Leu Pro Gln Met Pro Leu Phe Glu Asn Phe Tyr Ser Met Pro
 275 280 285
 Gln Gln Pro Ser Gln Gln Pro Gln Asp Phe Gly Leu Gln Pro Ala Gly
 290 295 300
 Pro Leu Gly Gln Ser His Leu Ala His His Ser Met Ala Pro Tyr Pro
 305 310 315 320
 Phe Pro Pro Asn Pro Asp Met Asn Pro Glu Leu Arg Lys Ala Leu Leu
 325 330 335
 Gln Asp Ser Ala Pro Gln Pro Ala Leu Pro Gln Val Gln Ile Pro Phe
 340 345 350
 Pro Arg Arg Ser Arg Arg Leu Ser Lys Glu Gly Ile Leu Pro Pro Ser
 355 360 365
 Ala Leu Asp Gly Ala Gly Thr Gln Pro Gly Gln Glu Ala Thr Gly Asn
 370 375 380
 Leu Phe Leu His His Trp Pro Leu Gln Gln Pro Pro Pro Gly Ser Leu
 385 390 395 400
 Gly Gln Pro His Pro Glu Ala Leu Gly Phe Pro Leu Glu Leu Arg Glu
 405 410 415
 Ser Gln Leu Leu Pro Asp Gly Glu Arg Leu Ala Pro Asn Gly Arg Glu
 420 425 430
 Arg Glu Ala Pro Ala Met Gly Ser Glu Glu Gly Met Arg Ala Val Ser
 435 440 445
 Thr Gly Asp Cys Gly Gln Val Leu Arg Gly Gly Val Ile Gln Ser Thr
 450 455 460
 Arg Arg Arg Arg Arg Ala Ser Gln Glu Ala Asn Leu Leu Thr Leu Ala
 465 470 475 480
 Gln Lys Ala Val Glu Leu Ala Ser Leu Gln Asn Ala Lys Asp Gly Ser
 485 490 495
 Gly Ser Glu Glu Lys Arg Lys Ser Val Leu Ala Ser Thr Thr Lys Cys
 500 505 510
 Gly Val Glu Phe Ser Glu Pro Ser Leu Ala Thr Lys Arg Ala Arg Glu
 515 520 525
 Asp Ser Gly Met Val Pro Leu Ile Ile Pro Val Ser Val Pro Val Arg
 530 535 540
 Thr Val Asp Pro Thr Glu Ala Ala Gln Ala Gly Gly Leu Asp Glu Asp
 545 550 555 560
 Gly Lys Gly Leu Glu Gln Asn Pro Ala Glu His Lys Pro Ser Val Ile
 565 570 575
 Val Thr Arg Arg Arg Ser Thr Arg Ile Pro Gly Thr Asp Ala Gln Ala
 580 585 590
 Gln Ala Glu Asp Met Asn Val Lys Leu Glu Gly Glu Pro Ser Val Arg

595					600					605					
Lys 610	Pro 610	Lys 610	Gln 610	Arg 610	Pro 615	Arg 615	Pro 615	Glu 615	Pro 615	Leu 615	Ile 620	Ile 620	Pro 620	Thr 620	Lys 620
Ala 625	Gly 625	Thr 625	Phe 625	Ile 630	Ala 630	Ser 635	Pro 635	Val 635	Tyr 635	Ser 635	Asn 635	Ile 635	Thr 635	Pro 640	Tyr 640
Gln 645	Ser 645	His 645	Leu 645	Arg 645	Ser 645	Pro 650	Val 650	Arg 650	Leu 650	Ala 650	Asp 655	His 655	Pro 655	Ser 655	Glu 655
Arg 660	Ser 660	Phe 660	Glu 660	Leu 660	Pro 660	Pro 665	Tyr 665	Thr 665	Pro 665	Pro 665	Pro 665	Ile 665	Leu 665	Ser 670	Pro 670
Val 675	Arg 675	Glu 675	Gly 675	Ser 675	Gly 675	Leu 680	Tyr 680	Phe 680	Asn 680	Ala 680	Ile 685	Ile 685	Ser 685	Thr 685	Ser 685
Thr 690	Ile 690	Pro 690	Ala 690	Pro 690	Pro 695	Pro 695	Ile 695	Thr 695	Pro 695	Lys 695	Ser 700	Ala 700	His 700	Arg 700	Thr 700
Leu 705	Leu 705	Arg 705	Thr 705	Asn 710	Ser 710	Ala 710	Glu 710	Val 710	Thr 710	Pro 715	Pro 715	Val 715	Leu 715	Ser 720	Val 720
Met 725	Gly 725	Glu 725	Ala 725	Thr 725	Pro 725	Val 725	Ser 725	Ile 725	Glu 730	Pro 730	Arg 730	Ile 730	Asn 735	Val 735	Gly 735
Ser 740	Arg 740	Phe 740	Gln 740	Ala 740	Glu 740	Ile 740	Pro 740	Leu 745	Met 745	Arg 745	Asp 745	Arg 745	Ala 750	Leu 750	Ala 750
Ala 755	Ala 755	Asp 755	Pro 755	His 755	Lys 755	Ala 760	Asp 760	Val 760	Trp 760	Gln 760	Pro 765	Trp 765	Glu 765	Asp 765	Ala 765
Leu 770	Glu 770	Ser 770	Ser 770	Arg 770	Glu 770	Lys 775	Gln 775	Arg 775	Gln 775	Val 775	Asp 775	Leu 775	Leu 775	Thr 775	Thr 775
Ala 785	Ala 785	Cys 785	Ser 785	Ser 785	Ile 785	Phe 790	Pro 790	Gly 790	Ala 790	Gly 795	Thr 795	Asn 795	Gln 795	Glu 795	Leu 795
Ala 805	Leu 805	His 805	Cys 805	Leu 805	His 805	Glu 805	Ser 805	Arg 805	Gly 810	Asp 810	Ile 810	Leu 810	Glu 810	Thr 815	Leu 815
Asn 820	Lys 820	Leu 820	Leu 820	Lys 820	Lys 820	Pro 820	Leu 825	Arg 825	Pro 825	His 825	Asn 825	His 825	Pro 825	Leu 825	Leu 825
Ala 835	Thr 835	Tyr 835	His 835	Tyr 835	Thr 835	Gly 840	Ser 840	Asp 840	Gln 840	Trp 840	Lys 840	Met 840	Ala 840	Glu 840	Arg 840
Lys 850	Leu 850	Phe 850	Asn 850	Lys 850	Gly 850	Ile 855	Ala 855	Ile 855	Tyr 855	Lys 855	Lys 855	Asp 855	Phe 855	Phe 855	Leu 855
Val 865	Gln 865	Lys 865	Leu 865	Ile 865	Gln 865	Thr 865	Lys 865	Thr 865	Val 865	Ala 865	Gln 865	Cys 865	Val 865	Glu 865	Phe 865
Tyr 875	Tyr 875	Thr 875	Tyr 875	Lys 875	Gln 875	Val 875	Lys 875	Ile 875	Gly 875	Arg 875	Asn 875	Gly 875	Thr 875	Leu 875	Leu 875
Thr 885	Phe 885	Gly 885	Asp 885	Val 885	Asp 885	Thr 885	Ser 885	Asp 885	Glu 885	Lys 885	Ser 885	Ala 885	Gln 885	Glu 885	Glu 885
Val 895	Glu 895	Val 895	Asp 895	Ile 895	Lys 895	Thr 895	Ser 895	Gln 895	Lys 895	Phe 895	Pro 895	Arg 895	Val 895	Pro 895	Leu 895
Pro 905	Arg 905	Arg 905	Glu 905	Ser 905	Pro 905	Ser 905	Glu 905	Glu 905	Arg 905	Leu 905	Glu 905	Pro 905	Lys 905	Arg 905	Glu 905
Val 915	Lys 915	Glu 915	Pro 915	Arg 915	Lys 915	Glu 915	Gly 915	Glu 915	Glu 915	Glu 915					

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 849 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

Ile	Arg	His	Glu	Val	Ser	Phe	Leu	Trp	Asn	Thr	Glu	Ala	Ala	Cys	Pro
1				5					10					15	
Ile	Gln	Thr	Thr	Asp	Thr	Asp	Gln	Ala	Cys	Ser	Ile	Arg	Asp	Pro	
			20				25					30			
Asn	Ser	Gly	Phe	Val	Phe	Asn	Leu	Asn	Pro	Leu	Asn	Ser	Ser	Gln	Gly
		35					40				45				

Tyr	Asn	Val	Ser	Gly	Ile	Gly	Lys	Ile	Phe	Met	Phe	Asn	Val	Cys	Gly
Thr	Met	Pro	Val	Cys	Gly	Thr	Ile	Leu	Gly	Lys	Pro	Ala	Ser	Gly	Cys
65					70					75					80
Glu	Ala	Glu	Thr	Gln	Thr	Glu	Glu	Leu	Lys	Asn	Trp	Lys	Pro	Ala	Arg
				85					90					95	
Pro	Val	Gly	Ile	Glu	Lys	Ser	Leu	Gln	Leu	Ser	Thr	Glu	Gly	Phe	Ile
			100					105					110		
Thr	Leu	Thr	Tyr	Lys	Gly	Pro	Leu	Ser	Ala	Lys	Gly	Thr	Ala	Asp	Ala
		115					120					125			
Phe	Ile	Val	Arg	Phe	Val	Cys	Asn	Asp	Asp	Val	Tyr	Ser	Gly	Pro	Leu
	130					135					140				
Lys	Phe	Leu	His	Gln	Asp	Ile	Asp	Ser	Gly	Gln	Gly	Ile	Arg	Asn	Thr
145					150					155					160
Tyr	Phe	Glu	Phe	Glu	Thr	Ala	Leu	Ala	Cys	Val	Pro	Ser	Pro	Val	Asp
				165					170					175	
Cys	Gln	Val	Thr	Asp	Leu	Ala	Gly	Asn	Glu	Tyr	Asp	Leu	Thr	Gly	Leu
			180					185					190		
Ser	Thr	Val	Arg	Lys	Pro	Trp	Thr	Ala	Val	Asp	Thr	Ser	Val	Asp	Gly
		195					200					205			
Arg	Lys	Arg	Thr	Phe	Tyr	Leu	Ser	Val	Cys	Asn	Pro	Leu	Pro	Tyr	Ile
					215						220				
Pro	Gly	Cys	Gln	Gly	Ser	Ala	Val	Gly	Ser	Cys	Leu	Val	Ser	Glu	Gly
225					230					235					240
Asn	Ser	Trp	Asn	Leu	Gly	Val	Val	Gln	Met	Ser	Pro	Gln	Ala	Ala	Ala
				245					250					255	
Asn	Gly	Ser	Leu	Ser	Ile	Met	Tyr	Val	Asn	Gly	Asp	Lys	Cys	Gly	Asn
			260					265					270		
Gln	Arg	Phe	Ser	Thr	Arg	Ile	Thr	Phe	Glu	Cys	Ala	Gln	Ile	Ser	Gly
		275					280					285			
Ser	Pro	Ala	Phe	Gln	Leu	Gln	Asp	Gly	Cys	Glu	Tyr	Val	Phe	Ile	Trp
		290				295					300				
Arg	Thr	Val	Glu	Ala	Cys	Pro	Val	Val	Arg	Val	Glu	Gly	Asp	Asn	Cys
305					310					315					320
Glu	Val	Lys	Asp	Pro	Arg	His	Gly	Asn	Leu	Tyr	Asp	Leu	Lys	Pro	Leu
				325					330					335	
Gly	Leu	Asn	Asp	Thr	Ile	Val	Ser	Ala	Gly	Glu	Tyr	Thr	Tyr	Tyr	Phe
			340					345					350		
Arg	Val	Cys	Gly	Lys	Leu	Ser	Ser	Asp	Val	Cys	Pro	Thr	Ser	Asp	Lys
		355					360					365			
Ser	Lys	Val	Val	Ser	Ser	Cys	Gln	Glu	Lys	Arg	Glu	Pro	Gln	Gly	Phe
					375						380				
His	Lys	Val	Ala	Gly	Leu	Leu	Thr	Gln	Lys	Leu	Thr	Tyr	Glu	Asn	Gly
385					390					395					400
Leu	Leu	Lys	Met	Asn	Phe	Thr	Gly	Gly	Asp	Thr	Cys	His	Lys	Val	Tyr
				40											

Cys Ser Leu Ala Asp Tyr Tyr Ile Arg Ser Gly His Phe Glu Lys Ala
 260 265 270
 Arg Asp Val Tyr Glu Glu Ala Ile Arg Thr Val Met Thr Val Arg Asp
 275 280 285
 Phe Thr Gln Val Phe Asp Ser Tyr Ala Gln Phe Glu Glu Ser Met Ile
 290 295 300
 Ala Ala Lys Met Glu Thr Ala Ser Glu Leu Gly Arg Glu Glu Glu Asp
 305 310 315 320
 Asp Val Asp Leu Glu Leu Arg Leu Ala Arg Phe Glu Gln Leu Ile Ser
 325 330 335
 Arg Arg Pro Leu Leu Leu Asn Ser Val Leu Leu Arg Gln Asn Pro His
 340 345 350
 His Val His Glu Trp His Lys Arg Val Ala Leu His Gln Gly Arg Pro
 355 360 365
 Arg Glu Ile Ile Asn Thr Tyr Thr Glu Ala Val Gln Thr Val Asp Pro
 370 375 380
 Phe Lys Ala Thr Gly Lys Pro His Thr Leu Trp Val Ala Phe Ala Lys
 385 390 395 400
 Phe Tyr Glu Asp Asn Gly Gln Leu Asp Asp Ala Arg Val Ile Leu Glu
 405 410 415
 Lys Ala Thr Lys Val Asn Phe Lys Gln Val Asp Asp Leu Ala Ser Val
 420 425 430
 Trp Cys Gln Cys Gly Glu Leu Glu Leu Arg His Glu Asn Tyr Asp Glu
 435 440 445
 Ala Leu Arg Leu Leu Arg Lys Ala Thr Ala Leu Pro Ala Arg Arg Ala
 450 455 460
 Glu Tyr Phe Asp Gly Ser Glu Pro Val Gln Asn Arg Val Tyr Lys Ser
 465 470 475 480
 Leu Lys Val Trp Ser Met Leu Ala Asp Leu Glu Glu Ser Leu Gly Thr
 485 490 495
 Phe Gln Ser Thr Lys Ala Val Tyr Asp Arg Ile Leu Asp Leu Arg Ile
 500 505 510
 Ala Thr Pro Gln Ile Val Ile Asn Tyr Ala Met Phe Leu Glu Glu His
 515 520 525
 Lys Tyr Phe Glu Glu Ser Phe Lys Ala Tyr Glu Arg Gly Ile Ser Leu
 530 535 540
 Phe Lys Trp Pro Asn Val Ser Asp Ile Trp Ser Thr Tyr Leu Thr Lys
 545 550 555 560
 Phe Ile Ala Arg Tyr Gly Gly Arg Lys Leu Glu Arg Ala Arg Asp Leu
 565 570 575
 Phe Glu Gln Ala Leu Asp Gly Cys Pro Pro Lys Tyr Ala Lys Thr Leu
 580 585 590
 Tyr Leu Leu Tyr Ala Gln Leu Glu Glu Glu Trp Gly Leu Ala Arg His
 595 600 605
 Ala Met Ala Val Tyr Glu Arg Ala Thr Arg Ala Val Glu Pro Ala Gln
 610 615 620
 Gln Tyr Asp Met Phe Asn Ile Tyr Ile Lys Arg Ala Ala Glu Ile Tyr
 625 630 635 640
 Gly Val Thr His Thr Arg Gly Ile Tyr Gln Lys Ala Ile Glu Val Leu
 645 650 655
 Ser Asp Glu His Ala Arg Glu Met Cys Leu Arg Phe Ala Asp Met Glu
 660 665 670
 Cys Lys Leu Gly Glu Ile Asp Arg Ala Arg Ala Ile Tyr Ser Phe Cys
 675 680 685
 Ser Gln Ile Cys Asp Pro Arg Thr Thr Gly Ala Phe Trp Gln Thr Trp
 690 695 700
 Lys Asp Phe Glu Val Arg His Gly Asn Glu Asp Thr Ile Lys Glu Met
 705 710 715 720
 Leu Arg Ile Arg Arg Ser Val Gln Ala Thr Tyr Asn Thr Gln Val Asn
 725 730 735
 Phe Met Ala Ser Gln Met Leu Lys Val Ser Gly Ser Ala Thr Gly Thr
 740 745 750
 Val Ser Asp Leu Ala Pro Gly Gln Ser Gly Met Asp Asp Met Lys Leu
 755 760 765
 Leu Glu Gln Arg Ala Glu Gln Leu Ala Ala Glu Ala Glu Arg Asp Gln
 770 775 780
 Pro Leu Arg Ala Gln Ser Lys Ile Leu Phe Val Arg Ser Asp Ala Ser
 785 790 795 800
 Arg Glu Glu Leu Ala Glu Leu Ala Gln Gln Val Asn Pro Glu Glu Ile
 805 810 815
 Gln Leu Gly Glu Asp Glu Asp Glu Met Asp Leu Glu Pro Asn
 820 825 830
 Glu Val Arg Leu Glu Gln Gln Ser Val Pro Ala Ala Val Phe Gly Ser

845

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 693 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

Met	Phe	Ser	Ala	Leu	Lys	Lys	Leu	Val	Gly	Ser	Asp	Gln	Ala	Pro	Gly
1				5				10						15	
Arg	Asp	Lys	Asn	Ile	Pro	Ala	Gly	Leu	Gln	Ser	Met	Asn	Gln	Ala	Leu
			20					25					30		
Gln	Arg	Arg	Phe	Ala	Lys	Gly	Val	Gln	Tyr	Asn	Met	Lys	Ile	Val	Ile
			35				40					45			
Arg	Gly	Asp	Arg	Asn	Thr	Gly	Lys	Thr	Ala	Leu	Trp	His	Arg	Leu	Gln
			50			55					60				
Gly	Arg	Pro	Phe	Val	Glu	Glu	Tyr	Ile	Pro	Thr	Gln	Glu	Ile	Gln	Val
65					70					75					80
Thr	Ser	Ile	His	Trp	Ser	Tyr	Lys	Thr	Asp	Asp	Ile	Val	Lys	Lys	Val
				85					90					95	
Glu	Val	Trp	Asp	Val	Val	Asp	Lys	Gly	Lys	Cys	Lys	Lys	Arg	Gly	Asp
			100					105					110		
Gly	Leu	Lys	Met	Glu	Asn	Asp	Pro	Gln	Glu	Xaa	Glu	Ser	Glu	Met	Ala
			115				120					125			
Leu	Asp	Ala	Glu	Phe	Leu	Asp	Val	Tyr	Lys	Asn	Cys	Asn	Gly	Val	Val
			130			135						140			
Met	Met	Phe	Asp	Ile	Thr	Lys	Gln	Trp	Thr	Phe	Asn	Tyr	Ile	Leu	Arg
145					150					155					160
Glu	Leu	Pro	Lys	Val	Pro	Thr	His	Val	Pro	Val	Cys	Val	Leu	Gly	Asn
				165					170					175	
Tyr	Arg	Asp	Met	Gly	Glu	His	Arg	Val	Ile	Leu	Pro	Asp	Asp	Val	Arg
			180					185					190		
Asp	Phe	Ile	Asp	Asn	Leu	Asp	Arg	Pro	Pro	Gly	Ser	Ser	Tyr	Phe	Arg
			195				200					205			
Tyr	Ala	Glu	Ser	Ser	Met	Lys	Asn	Ser	Phe	Gly	Leu	Lys	Tyr	Leu	His
						215					220				
Lys	Phe	Phe	Asn	Ile	Pro	Phe	Leu	Gln	Leu	Gln	Arg	Glu	Thr	Leu	Leu
225					230					235					240
Arg	Gln	Leu	Glu	Thr	Asn	Gln	Leu	Asp	Met	Asp	Ala	Thr	Leu	Glu	Glu
				245					250					255	
Leu	Ser	Val	Gln	Glu	Glu	Thr	Glu	Asp	Gln	Asn	Tyr	Gly	Ile	Phe	Leu
			260					265					270		
Glu	Met	Met	Glu	Ala	Arg	Ser	Arg	Gly	His	Ala	Ser	Pro	Leu	Ala	Ala
		275					280					285			
Asn	Gly	Gln	Ser	Pro	Ser	Pro	Gly	Ser	Gln	Ser	Pro	Val	Leu	Pro	Ala
						295					300				
Pro	Ala	Val	Ser	Thr	Gly	Ser	Ser	Ser	Pro	Gly	Thr	Pro	Gln	Pro	Ala
305					310					315					320
Pro	Gln	Leu	Pro	Leu	Asn	Ala	Ala	Pro	Pro	Ser	Ser	Val	Pro	Pro	Val
				325					330					335	
Pro	Pro	Ser	Glu	Ala	Leu	Pro	Pro	Pro	Ala	Cys	Pro	Ser	Ala	Pro	Ala
			340					345					350		
Pro	Arg	Arg	Ser	Ile	Ile	Ser	Arg								

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

00712813.11300

Gln Asp Ile Thr Leu Ser Ser Glu Glu Glu Ala Glu Val Ala Ala Pro
 465 470 475 480
 Thr Lys Gly Pro Ala Pro Ala Pro Gln Gln Cys Ser Glu Pro Glu Thr
 485 490 495
 Lys Trp Ser Ser Ile Pro Ala Ser Lys Pro Arg Arg Gly Thr Ala Pro
 500 505 510
 Thr Arg Thr Ala Ala Pro Pro Trp Pro Gly Gly Val Ser Val Arg Thr
 515 520 525
 Gly Pro Glu Lys Arg Ser Ser Thr Arg Pro Pro Ala Glu Met Glu Pro
 530 535 540
 Gly Lys Gly Glu Gln Ala Ser Ser Ser Glu Ser Asp Pro Glu Gly Pro
 545 550 555 560
 Ile Ala Ala Gln Met Leu Ser Phe Val Met Asp Asp Pro Asp Phe Glu
 565 570 575
 Ser Glu Gly Ser Asp Thr Gln Arg Arg Ala Asp Asp Phe Pro Val Arg
 580 585 590
 Asp Asp Pro Ser Asp Val Thr Asp Glu Asp Glu Gly Pro Ala Glu Pro
 595 600 605
 Pro Pro Pro Pro Lys Leu Pro Leu Pro Ala Phe Arg Leu Lys Asn Asp
 610 615 620
 Ser Asp Leu Phe Gly Leu Gly Leu Glu Glu Ala Gly Pro Lys Glu Ser
 625 630 635 640
 Ser Glu Glu Gly Lys Glu Gly Lys Thr Pro Ser Lys Glu Lys Lys Lys
 645 650 655
 Lys Thr Lys Ser Phe Ser Arg Val Leu Leu Glu Arg Pro Arg Ala His
 660 665 670
 Arg Phe Ser Thr Arg Val Gly Tyr Gln Val Ser Val Pro Asn Ser Pro
 675 680 685
 Tyr Ser Glu Ser Tyr
 690